UNIVERSITÀ DEGLI STUDI FIRENZE

DOTTORATO DI RICERCA
INTERNATIONAL DOCTORATE IN STRUCTURAL BIOLOGY

CICLO XXXII

COORDINATOR Prof. Claudio Luchinat

Bioinformatics tools for metalloprotein analysis

Settore Scientifico Disciplinare CHIM/03

PhD student
Dott. Valeria Putignano

Tutor
Prof. Claudia Andreini

Coordinator
Prof. Claudio Luchinat

November 2016 – 2019

This thesis has been approved by the University of Florence, the University of Frankfurt and the Utrecht University

Universiteit Utrecht

CERM Firenze

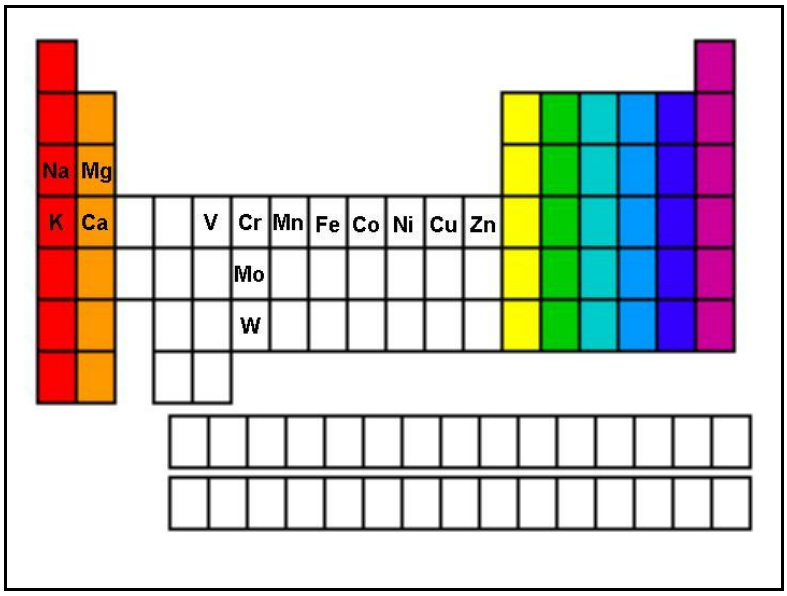
1 INTRODUCTION

1.1 The importance of metals in biology

With the advent of the so-called *bioinorganic chemistry* (the discipline at the interface of chemistry and biology) since the 70's and its rapid development during the past years, the significant role of metal ions in biological systems, including their interplay with proteins, has become evident [1].

Metal elements are classified in respect to their biological behavior into two different classes: *essential trace elements* (Figure 1), which are indispensable for normal life of the organisms [2], and *toxic elements*, whose assimilation may determine the alteration of cell functioning and eventually be lethal to the organism [3].

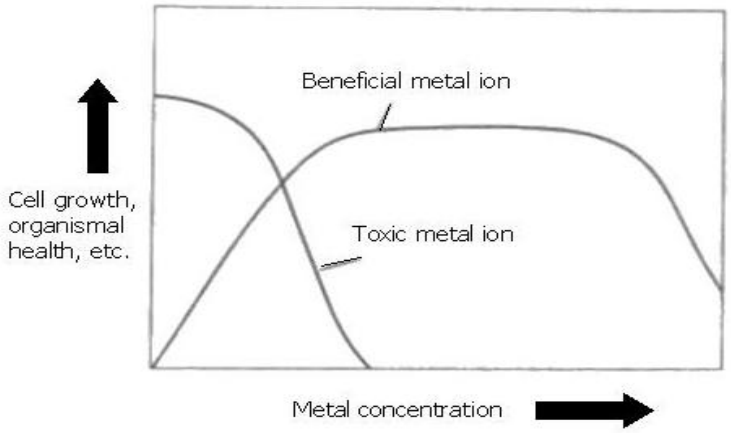
Figure 1. *Simplified version of the periodic table showing important elements in bioinorganic chemistry ('essential trace elements')*



Although the term 'toxic' is usually referred to certain metals such as mercury, aluminum and cadmium, it can be applied to all metal ions because all elements are toxic if they are present in living organisms in sufficiently high quantity. As Paracelsius stated almost 500 years ago, "it is the dose that makes it poison or remedy". Obviously, the dose at which specific metals become toxic varies greatly [4]. Indeed, for any element a curve as that shown in Figure 2 exists, in which the physiological response of the organism is reported as a function of the

assumed quantity. If the concentration of a given essential metal is too low, processes that need to use that ion will be adversely affected and the organism will suffer from metal ion deficiency [5]. Once the concentration of a given metal ion is above a lower threshold, there will be enough of that ion to fulfil its biological functions. However, the concentration cannot be increased indefinitely without adverse consequences. Above an upper threshold, the effects of metal ion toxicity will arise [6]. For example, a metal ion might bind to an inappropriate site, competing with other beneficial metal ions for that site; furthermore, there might be undesirable reactivity of the metal ion when it is not properly controlled in its normal binding sites. These effects can be illustrated with the example of iron, which is an essential ion for all organisms, including human. Iron is involved, among other functions, in dioxygen transport and in a variety of electron-transport pathways. Iron deficiency resulting from diminished supply or uptake of the metal, or from its loss, produces anemia because of inadequate quantities required for hemoglobin synthesis [7]. On the other hand, iron overload can also occur, e.g., by accidental ingestion: the excess iron can accumulate and is not easily excreted. Once the iron-storage mechanisms are saturated, excess amounts of the metal are released into the cell, where they can catalyse the formation of various oxygen-based free radicals and extensively damage tissues [8].

Figure 2. *Representation of the concentration dependence of the toxic and beneficial effects of metal ions*



Some metals have no known or presumed biological function: when present in cells, they may be rather innocuous or quite toxic. The preminent factor in determining the appearance of Figure 2, thus the biological behavior of each metal in living organisms, is the environment in which life first began and then evolved. As a result of the evolutionary

process, iron, zinc, copper, magnesium, manganese and other metal ions are crucial to life today [9].

1.2 Metals in cells

As mentioned in the previous section, the concentration of metal ions in cells must be maintained within proper ranges [10]. *Homeostasis*, the maintenance of the concentration of beneficial metal ions in the correct range, and *detoxification*, the removal of toxic concentrations of non-beneficial metal ions, require balance between the processes of metal ion uptake, utilization, storage and excretion.

Bioaccumulation of metals in cells reflects a number of factors: (i) ecological (e.g., close contact with environment); (ii) physiological (e.g., filtering activity to satisfy respiratory and nutritional needs); and (iii) biochemical (e.g., metal tolerance strategies that involve metal sequestration, inclusion or elimination). Cells actively maintain relatively high intracellular concentrations of the essential metal ions: for instance, some studies of transition metal quotas in *Escherichia coli* reveal that individual bacteria concentrate Zn and Fe by several orders of magnitude relative to the concentration in a typical growth medium [11]. On the other hand, concentration of the free forms of Zn ions within the cytoplasm are proposed to be lower than 0.5 fM, which is an extraordinarily low threshold [12]. These observations lead to the conclusion that if the transition metals are abundant in the cell, then also metal-binding proteins must be: as a matter of fact, metal-binding proteins correspond to about thirty pecent of all protein structures contained in the Protein Data Bank[31]. But, how do cells allocate the correct metals to specific protein sites, while avoiding toxic side reactions at such high total concentrations of metal ions? A mechanism for this process appears to be the evolution of specific pathways involving several proteins (transporters, metallochaperones) which protect and drive the metal ions through the cytoplasm, ultimately transferring them to specific target proteins [13].

Metal ions not utilized in biological systems can be quite toxic, often because they tend to bind non-specifically, but with high affinity, to certain types of sites. Because of this tight binding, which is often a consequence of kinetic inertness, these metals may bind to sites where they inhibit some normal processes in such a manner that they are not easily removed and excreted. Other possible causes of metal ion toxicity include the formation of insoluble salts in biological fluids, participation in hydrolytic reactions that degrade biopolymers, or redox chemistry that produces damaging by-products, such as hydroxyl radicals. For these

reasons, some metal ions are toxic to cells at all concentrations, therefore detoxification systems that employ a variety of mechanisms to rid the cell of these potentially lethal toxins have evolved [14]. In most bacterial organisms, the expression of metal resistance systems is controlled at the level of transcription by sensor proteins that 'sense' specific metal ions via their direct coordination [15].

In conclusion, cells can manage metal-protein speciation: they acquire more of those ions which are deficient, while exporting or sequestering those that are in surplus or toxic. The beneficial intracellular concentration of metals is maintained by the strict regulation in the expression of proteins involved in specific metal uptake, export or storage.

1.3 Biological roles of metals

It is well known that metal-binding proteins participate in some of the most important biochemical processes including respiration, most of metabolic processes, nitrogen fixation, photosynthesis, development, signal transduction and many others. In all of these proteins the first coordination sphere of each metal ion is in general referred to as the *metal site*, which can be classified into four basic types depending on the function:

(i) *Structural*: when the metal stabilizes the tertiary or quaternary structure of the protein and/or modulates the interaction of the protein with the substrate/protein target (e.g., zinc-fingers).

(ii) *Catalytic*: when the biochemical enviroment created by the coordination of the ion and the global structure of the protein modulates biochemical properties (charge distribution, protein stability, redox potential, etc.) determining the conditions of reactions. In other words, the bound metal is mandatory for the protein to carry out its physiological function (e.g., carbonic anhydrase). It is worth to notice that metal ions are found to be bound to all the six classes of enzymes defined by the International Union of Biochemistry and Molecular Biology.

(iii) *Dioxygen transport*: when the metal binds/release O_2 in respiration (e.g., hemoglobin).

(iv) *Electron transport*: when the metal in the protein undergoes redox reaction without themselves catalysing an overall chemical change in a substrate molecule (e.g., cytochrome *c*).

(v) *Storage*: when the metal is bound to a protein involved in the homeostasis of the ion. These proteins have the function to uptake, hold and release the metals in response to the cell demand (e.g., metallothionein).

The same metal can play different roles depending on its chemical context in the macromolecular environment. However the functions that an ion can perform in proteins is intimately linked to the physico-chemical properties of the element (redox properties, Lewis acidity, etc.). The non-redox ions, such as Zn^{2+} often are bound to proteins to confer them stability; further, Zn^{2+} is an effective Lewis acid catalyst in a wide range of transformations not involving electron transfer. Electron transfer and redox centres generally occur at sites containing iron or copper, also molybdenum and tungsten catalyse oxidation-reduction reactions. Divalent nickel is a Lewis acid catalyst (e.g., urease) but is also involved in enzymes where redox activity is required (e.g., [Ni-Fe]- hydrogenases, carbon monoxide dehydrogenase). Magnesium normally exhibits a structural and certain catalytic functions (e.g., ATPase). Calcium also functions as a structural metal site and acts as a trigger in intracellular messenger systems controlling processes such as muscle contraction, secretion, glycolysis and ion transport.

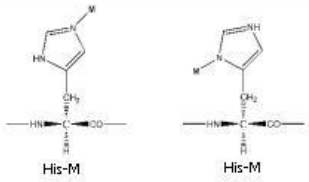
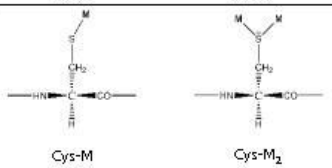
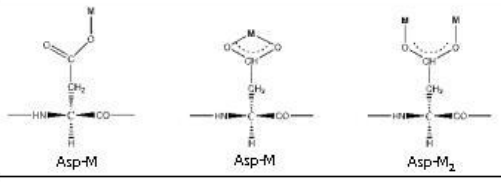
In agreement with the importance of the roles covered by metals and by metal-binding proteins associated with them, in the last few years it has become increasingly clear that several pathologies are associated with malfunction of the metabolism of metal-containing systems. Dysregulation of metal homeostasis may be involved in carcinogenesis as well as in metastasis formation and progress, and has been associated with cardiovascular diseases and several neurodegenerative disorders such as ALS, Menkes, Wilson's, Alzheimer's and Parkinson's diseases [16,17]. Furthermore, metal ions and metal-binding proteins play crucial roles in determining bacterial virulence, as well as in the development of antibiotic resistance by pathogenic microorganisms [18]. This scenario placed the study of metal-binding proteins at the forefront not only in bioinorganic chemistry, the field of science that studies the interplay between metal ions and biological systems, but also in the biomedical and drug discovery research.

1.4 How proteins bind metals

From the point of view of metal coordination, a polypeptide chain can be regarded as a polydentate ligand. Metals usually are bound to the polypeptide through nitrogen, oxygen and sulfur provided by *endogenous* ligands [19]. The amino acids that commonly function as ligands

and their modes of interaction are shown in Figure 3. The most common side-chain ligands are the thiolate group of cysteine, the imidazole group of histidine, the carboxylate group of glutamic and aspartic acids, and the phenolate group of tyrosine. With the exception of tyrosine, each of these residues has been observed in a few cases to act as a bridging ligand between two metal ions and to serve as a terminal ligand to a single ion. Less frequently encountered metal donors are the hydroxyl groups of serine and threonine, the thioether group of methionine, the carboxamide groups of glutamine and asparagines and the amino group of lysine. In addition to the donor atoms provided by side-chains, metal ions can also bind to backbone carbonyl groups, deprotonated backbone nitrogen atoms and the N-terminal amino and C-terminal carboxyl groups. Protic acids coordinate as anions; from the tabulated pK_a values (Figure 3), only carboxylate is available in a substantially deprotonated form around neutral pH. However, these values are generally expected to may vary by about 1 log unit in proteins, owing to dielectric and local electrostatic effects. Metals can bind ligands at pH values well below their pK_a's. As an example, coordination of a metal ion at the unprotoned nitrogen atom of the imidazolyl group lower the pK_a of the protonated nitrogen by about 2 log units due to an inductive effect. The ability of a metal to compete effectively with a proton in ligand binding is dictated in large measure by the strength of the metal-ligand bond.

Figure 3. *Most common endogenous biological ligands and their approximate pKa values*

Residue	Complexes	pKa
His		~6.0
Cys		~8.3
Asp/Glu	 Similar structures are formed by glutamic acid	~3.6 / ~4.25
Tyr		~10.1

Ligands not derived from proteins are considered *exogenous.* Water is the most frequent exogenous ligand. Coordination of water results in a substantial lowering of its pK_a value because the inductive effect of a bound cation further polarizes the O-H bond. This effect increases as the effective nuclear charge of the metal ion increases and its radius decreases.

The nature of the metal ion and its physico-chemical properties determine coordination preferences which influence the capability of proteins to discriminate among metals in cell and use them to carry out their physiological function [20]: evolution "knows" and "uses" these preferences to create molecules more and more selective and/or "intelligent". Metal ions generally bind to donor ligands according to preferences dictated by the hard-soft theory of acids and bases as reported in Table 1. So, alkaly and alkaline metals (i.e. Ca^{2+}) are most often coordinated in proteins by carboxylate groups (e.g., Asp, Glu) whereas for instance Cu^{+} prefers soft donors such RS^{-} ligands in cysteinyl side-chains. Border-line ions generally show a larger variety in coordinating ligands, although they are prevalently bound to nitrogen donors. Also the geometry coordination preferences are important: in protein sites, ligands are often arranged in the three-dimensional (3D) space according to the metal preferences, in particular when the metal must be bound strictly by the protein [21]. In fact, alterations in the ligand donor atoms and in the stereochemistry at the metal centre can dramatically change the relative metal affinities of the site, as well as some properties of the bound metal such as acid-base reactivity and redox potential. In living systems, the metal does not always need to be tightly bound: proteins often use low-affinity sites and finely tune the features of metal coordination to carry out particular functions. *In vitro* experiments have shown that there exist proteins which can bind different ions with different geometry coordination preferences at the same regulatory metal site: one activates the protein, whereas the other inhibits it [22].

Although the properties of a metal center in a biological enviroment are primarily determined by the first coordination sphere of the metal, also residues which are not directly coordinating may contribute to increase/reduce the thermodynamic stability of the site. Such residues can influence the local hydrophilicity/hydrophobicity, cause the steric blockage of the coordination sites, and provide hydrogen-bonding groups that can interact with bonded and non-bonded atoms in the coordination sphere of the metal.

Finally, it has to be noted that proteins not always can discriminate different metal ions only on the basis of the coordination chemistry, so in many cases molecular recognition

occurs through metal partitioning in the cell: some cellular pathways evolved with the only task to locate the fair metal to the fair protein.

Table 1: *Some biologically essential metal ions and their correspondent common oxidation states and the conseguent external electronical configuration, their common coordination numbers* [23].

Metal	Common oxidation states	d^n	Hard/soft properties	Common coordination number
Fe	+2	d^6	Borderline	4-5-**6**
Fe	+3	d^5	Hard	4-5-**6**
Zn	+2	d^{10}	Borderline	**4**-5-6
Cu	+1	d^{10}	Soft	2-3-**4**
Cu	+2	d^9	Borderline	**4**-5-6

2 STATE OF THE ART

2.1 MetalPDB: a central web resource for metal-binding proteins

With the aim of providing the scientific community with tools for the analysis of biomolecules, bioinformatics, i.e. the discipline applying informatics to the study of biological systems, has made available plenty of databases and predictive software. Nevertheless, very few of these resources have been dedicated to the study of metal-binding proteins (Table 2), probably because metals confer to biomolecules properties that are peculiar and difficult to encode. The first attempts of collecting and organizing all the available information on metal-binding proteins into databases date back to the end of 90s, and include, for example, PROMISE [24] and MDB [25].

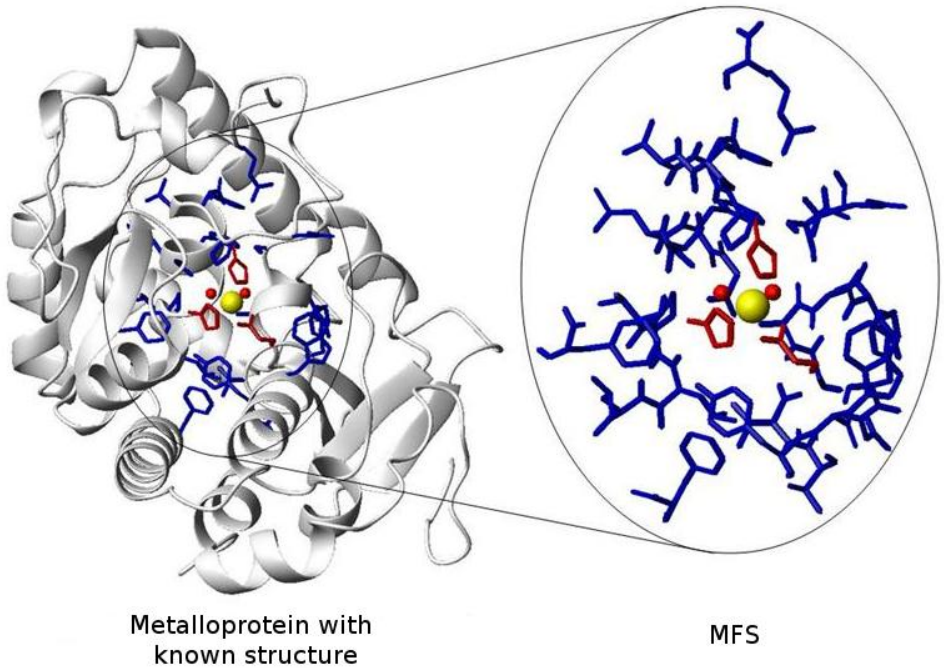
Table 2: *Resources dedicated to the study of metal-binding proteins*

The **MDB** [25] (http://metallo.scripps.edu/) is the first database that was created for metal-binding proteins and is specifically geared toward providing information useful for metal-binding protein design. This results in the information provided consisting mainly of a description of the features of the metal coordination environment. This database has not been updated since 2003.
COMe [26] (http://www.flymine.org/come/) provides only information on the first coordination sphere of the metal center, i.e. essentially what MetalPDB is providing in the first coordination sphere tab. This database has not been updated since 2005.
MESPEUS [27] (http://mespeus.bch.ed.ac.uk/MESPEUS_10/) is a relatively recent database, implemented in 2008, which provides extensive information on the metal coordination environment of metal-binding proteins in the PDB [28], basically providing a detailed description of all geometric features of the metal site. Crystallographic features are also described extensively, and it is possible to easily generate statistics for metals in any selected environment. Whereas MESPEUS geometric insight is far more extended than what we are providing in MetalPDB, its usefulness for functional analysis is more limited. Indeed, MESPEUS does not provide any comparison between different sites, as we instead accomplish by looking at equivalent and equistructural sites, nor it provides any analysis of protein domains. This database has not been updated since 2010.
MetLigDB [29] (http://silver.sejong.ac.kr/MetLigDB/home.html) focuses on the analysis of organic ligands binding to metal-binding proteins and not on metal–biomacromolecule interactions; its scope is thus widely different than MetalPDB.

MINAS [30] (http://www.minas.uzh.ch/) focuses on metal–nucleic acid interactions, and thus it does not include metal-binding proteins. Thus MetalPDB and MINAS can be seen as complementary, with some limited overlap. Note that of the 175 115 MFS contained in MetalPDB, 86 637 (49.5%) have at least one protein ligand and no nucleic acid ligand whereas 31 452 (18.0%) have at least one nucleic acid ligand and no protein ligand and 54 594 (31.2%) have ligands that are neither proteic or nucleic (the latter MFS's may however interact with proteins and/or nucleic acids in their second sphere).

Currently, the most exhaustive collection of data relevant to metal-binding proteins is MetalPDB [31] (http://metalweb.cerm.unifi.it), a resource developed by our group. The information in MetalPDB derives from the automated analysis of all the 3D structures of the adducts between biological macromolecules and metal ions or metal-containing cofactors available from the Protein Data Bank [28] (PDB). The central objects of MetalPDB are the Minimal Functional Sites (MFSs), which are 3D templates that describe the local environment around the metal(s) independently of the larger context of the macromolecular structure embedding the site(s). In particular, MFSs comprise the metal ion, its ligands and any chemical species within 5 Å from a ligand (Figure 4).

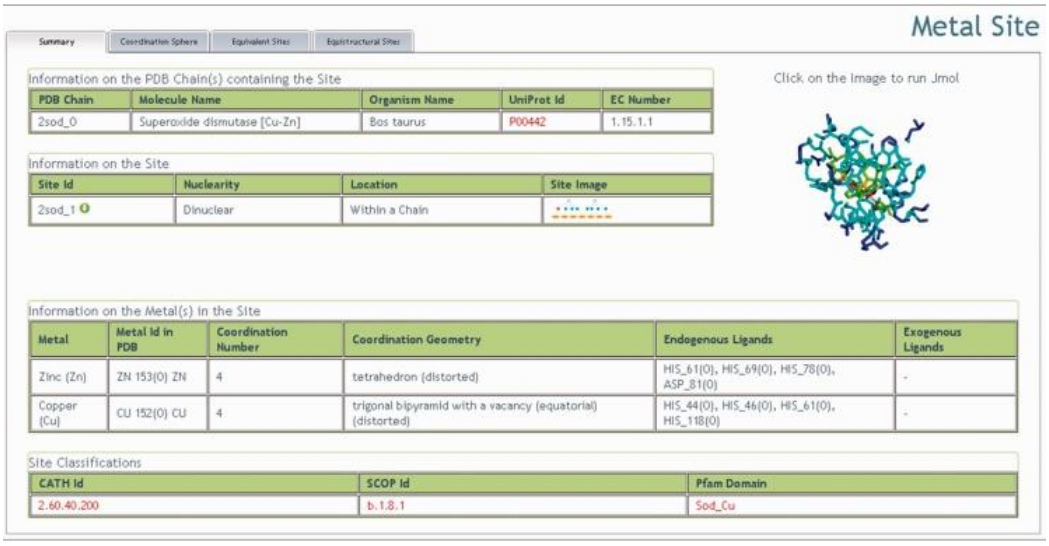
Figure 4. *Example of minimal functional site (MFS)*



Such 3D models have several advantages: they can be straightforwardly extracted from PDB structures, can be automatically compared via structural alignment to generate classifications, and, most importantly, embed the information on the chemico-physical determinants of the properties of the site, and thus of the metal function. It is well established, indeed, that the local environment of the metal ion also beyond its ligands (e.g. H-bonds, salt-bridges between ligands and neighboring atoms) have an important role in tuning its chemical reactivity.

MetalPDB allows users to query data using a web interface available at http://metalweb.cerm.unifi.it/. Searches may return a single database entry (e.g. when searching by PDB code) or multiple entries (e.g. sequence searches). In the first release of the resource the information of each site was organized into four different pages: a *Summary page, a Coordination sphere page, an Equistructural sites page* and *an Equivalent sites page*. The page shown by default is the *Summary page* (Figure 5) , including general information such as the EC number of the amino acid chain(s) containing the site, the coordination geometry of each metal in the site [32], and the structural or domain classification of the chain(s) containing the site.

Figure 5. *MetalPDB summary page for carbonic anhydrase 2 (12CA).*



The *Coordination sphere* page (Figure 6) provides more detailed information for each metal in the site on coordination as well as other structural properties. Indeed, the tab contains a large table for each metal that is further subdivided to display or permit access to metal properties. For example, donor atom names, types and distances from the metal are given in

tabular form. In addition, for each ligand it is possible to display and/or download tables reporting hydrogen bonding or van der Waals interactions. The same information can be schematically visualized. The rightmost column of each metal table shows a plot of the metal environment.

Figure 6. *MetalPDB coordination sphere page for carbonic anhydrase 2 (12CA).*

Under the *Equivalent sites* page (Figure 7) the user can find a list of sites that are equivalent to the site currently displayed (see the 'Database construction' section). Equivalent sites can be found in different PDB structures having the same fold, or in different but identically folded chains within the same PDB structure. In a nutshell, the list of equivalent sites contains all MFSs present in the PDB databank that contain the same metal in the same position as the current MFS, within a structure with the same fold as the structure containing the current site. However, the ligands may differ, although this is not common. Instead, the neighbors to the ligands will typically differ, to an extent depending largely on the sequence similarity between the protein chains compared [33]. Thus, the Equivalent sites tab allows users to readily identify families of proteins containing the same MFS, facilitating them to deal with the far from trivial task of assessing the redundancy of PDB structures in terms of their metal content. The coordinates of all the superimposed sites can be immediately downloaded from MetalPDB, together with a very simple Pymol (https://pymol.org/2/) script to visualize them.

Figure 7. *MetalPDB equivalent sites page for carbonic anhydrase 2 (12CA).*

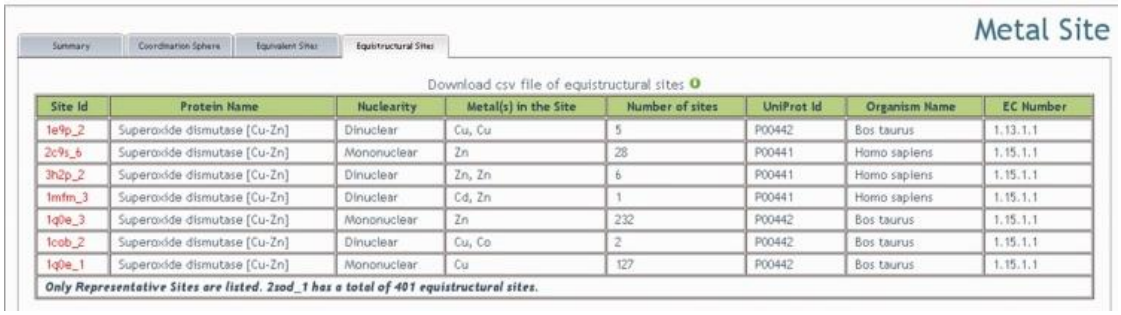
Download equivalent sites Download csv file of equivalent sites

Site Id	Protein Name	UniProt Id	Organism Name	EC Number	Metal(s) - Proteic Metal-binding Pattern
1kwq_1	Carbonic anhydrase 2	P00918	Homo sapiens	4.2.1.1	Zn_262 - p1: HX(1)HX(22)H
1kwr_1	Carbonic anhydrase 2	P00918	Homo sapiens	4.2.1.1	Zn_262 - p1: HX(1)HX(22)H
3lxe_1	Carbonic anhydrase 1	P00915	Homo sapiens	4.2.1.1	Zn_261 - p1: HX(1)HX(22)H
3f4x_1	Carbonic anhydrase 2	P00918	Homo sapiens	4.2.1.1	Zn_262 - p1: HX(1)HX(22)H
3f7u_1	Carbonic anhydrase 4	P22748	Homo sapiens	4.2.1.1	Zn_260 - p1: HX(1)HX(22)H
3f7u_2	Carbonic anhydrase 4	P22748	Homo sapiens	4.2.1.1	Zn_263 - p1: HX(1)HX(22)H
3s71_1	Carbonic anhydrase 2	P00918	Homo sapiens	4.2.1.1	Zn_262 - p1: HX(1)HX(22)H
3s76_1	Carbonic anhydrase 2	P00918	Homo sapiens	4.2.1.1	Zn_1 - p1: HX(1)HX(22)H
3s75_1	Carbonic anhydrase 2	P00918	Homo sapiens	4.2.1.1	Zn_262 - p1: HX(1)HX(22)H
1ydb_1	Carbonic anhydrase 2	P00918	Homo sapiens	4.2.1.1	Zn_262 - p1: HX(1)HX(22)H
1ydc_1	Carbonic anhydrase 2	P00918	Homo sapiens	4.2.1.1	Zn_262 - p1: HX(1)HX(22)H
4uov_4	Carbonate dehydratase	E8T502	Thermovibrio ammonificans	4.2.1.1	Zn_298 - p1: HX(1)HX(16)H
2weh_1	Carbonic anhydrase 2	P00918	Homo sapiens	4.2.1.1	Zn_1262 - p1: HX(1)HX(22)H
5sz3_1	Carbonic anhydrase 2	P00918	Homo sapiens	4.2.1.1	Zn_301 - p1: HX(1)HX(22)H
5flo_1	Carbonic anhydrase 2	P00918	Homo sapiens	4.2.1.1	Zn_1262 - p1: HX(1)HX(22)H
5fnh_1	Carbonic anhydrase 2	P00918	Homo sapiens	4.2.1.1	Zn_1262 - p1: HX(1)HX(22)H
4n0x_1	Carbonic anhydrase 2	P00918	Homo sapiens	4.2.1.1	Zn_301 - p1: HX(1)HX(22)H
4qj0_3	Carbonic anhydrase 12	O43570	Homo sapiens	4.2.1.1	Zn_301 - p1: HX(1)HX(23)H
4qjw_4	Carbonic anhydrase 12	O43570	Homo sapiens	4.2.1.1	Zn_301 - p1: HX(1)HX(23)H
4q8x_2	Carbonic anhydrase 2	P00918	Homo sapiens	4.2.1.1	Zn_301 - p1: HX(1)HX(22)H
4rn4_1	Carbonic anhydrase 2	P00918	Homo sapiens	4.2.1.1	Zn_301 - p1: HX(1)HX(22)H

Under the *Equistructural sites* tab (Figure 8) the user can find a list of sites that are equistructural to the MFS currently displayed (see the 'Database construction' section). As previously noted, two equistructural sites may or may not be also equivalent. For simplicity, the download button in the MetalPDB interface allows users to download a table of equistructural sites that are not equivalent to the MFS of interest (the latter can be obtained via the Equivalent sites tab). In practice, MFSs that are equistructural but not equivalent are sites in corresponding positions within protein structures having the same fold while they differ for their metal contents. This can happen for a variety of reasons. Metal ions can replace one another within the same site for both physiological and non-physiological reasons [34,35] or upon *in vitro* chemical treatment [typically to introduce spectroscopically active metals [36,37]]. Engineering of the metal ligands or of their neighbors can affect the relative affinity of a site toward different metal ions, eventually leading to incorporation of different metals in mutants with respect to the wild-type protein [38,39]. For polynuclear sites, it is additionally possible to observe phenomena such as the incorporation of different sets of metal ions (which again can be physiologically relevant or entirely due to *in vitro* treatment, and can change the nuclearity of the site), replacement of some or all of the metal ions with others [e.g. as observed in phosphatases [40,41]]. Each equistructural site shown in the tab is the representative (i.e. the site in the structure with the highest resolution) of a group of equivalent MFSs: the sites equivalent to these representatives are not shown to allow users to

grasp immediately the variation range independently of the number of MFSs in each group (Figure 8).

Figure 8. *MetalPDB equistructural sites page for carbonic anhydrase 2 (12CA).*



Site Id	Protein Name	Nuclearity	Metal(s) in the Site	Number of sites	UniProt Id	Organism Name	EC Number
1e9p_2	Superoxide dismutase [Cu-Zn]	Dinuclear	Cu, Cu	5	P00442	Bos taurus	1.13.1.1
2c9s_6	Superoxide dismutase [Cu-Zn]	Mononuclear	Zn	28	P00441	Homo sapiens	1.15.1.1
3h2p_2	Superoxide dismutase [Cu-Zn]	Dinuclear	Zn, Zn	6	P00441	Homo sapiens	1.15.1.1
1mfm_3	Superoxide dismutase [Cu-Zn]	Dinuclear	Cd, Zn	1	P00441	Homo sapiens	1.15.1.1
1q0e_3	Superoxide dismutase [Cu-Zn]	Mononuclear	Zn	232	P00442	Bos taurus	1.15.1.1
1cob_2	Superoxide dismutase [Cu-Zn]	Dinuclear	Cu, Co	2	P00442	Bos taurus	1.15.1.1
1q0e_1	Superoxide dismutase [Cu-Zn]	Mononuclear	Cu	127	P00442	Bos taurus	1.15.1.1
Only Representative Sites are listed. 2sod_1 has a total of 401 equistructural sites.							

MetalPDB is updated periodically in an automated manner, as described in Table 3.

Table 3: *Pipeline of MetalPDB update*

1	Download the coordinates for all structures in the PDB.
2	Process each coordinate file to identify all metal atoms in the structure.
3	For each metal atom in each structure from step [42] identify the ligands to it. Ligands are chemical species that contain at least one non-hydrogen atom at a distance smaller than 3.0 Å from the metal. They can be residues in a polypeptide or a polynucleotide chain (endogenous ligands) as well as different ions or molecules such as water, sulfide, acetate (exogenous ligands). Organic cofactors such as heme are considered exogenous ligands.
4	Each pair of metal atoms that have at least one common ligand, such as a bridging amino acidic side chain or exogenous anion, or whose distance is lower than 5 Å is included into a single polynuclear site. This procedure is iterated such that if metal A and metal B are to be included into a single site and then metal B and metal C are also to be included in a single site, eventually a three-nuclear site is formed that contains all three metal ions. This procedure allowed us to define, e.g., each Fe_4S_4 cluster found in ferredoxins as an individual four-nuclear site.

5	Identify the neighbors of all the ligands (both endogenous and exogenous) to the metal atom(s) in each mono- or polynuclear site. Ligand neighbors are chemical species (residues in a polypeptide or a polynucleotide chain, or other molecules or ions) that contain at least one non-hydrogen atom at a distance smaller than 5.0 Å from the ligand itself. The ensemble of the neighbors, the ligands and the metal atom(s) constitute the MFS. H-bond interactions between ligands and ligand neighbors are identified using the HBPLUS program [43].
6	For each protein chain in a PDB structure, identify the 50% sequence identity group in the PDB, the EC number, if relevant, as well as the UniProt[44] (http://www.uniprot.org/), CATH[45] (http://www.cathdb.info/), SCOP[46] (http://scop.mrc-lmb.cam.ac.uk/scop/) and Pfam[47] (http://pfam.sanger.ac.uk/) codes. Each MFS is then associated with the CATH, SCOP and Pfam code(s) of the protein domain(s) that contain the ligands.
7	Group MFSs into sets of 'equivalent' and 'equistructural' MFSs. Two MFSs are defined to be 'equivalent' when they satisfy the following conditions: (i) they have the same CATH, SCOP or Pfam classification; alternatively, the sequence identity between the two PDB chains that contain them is ≥50% (effectively meaning that the two chains have the same fold (19)); (ii) after structural superposition of the PDB chains containing them, the two MFSs are superimposed (i.e. the distance between their geometric centers is <3.5 Å); and (iii) after structural superposition of the PDB chains containing them, the two MFSs have the same metal elements in the same positions. For the latter condition to be fulfilled, equivalent sites must have the same nuclearity. Two MFSs are defined to be 'equistructural' when they satisfy conditions (i) and (ii) above, while condition (iii) does not need to be fulfilled. This implies that two equivalent sites are also equistructural, but the converse is not necessarily true. All equivalent and equistructural MFSs are grouped into clusters of equivalent and equistructural MFSs, respectively, by using a single linkage clustering strategy. For each group of equivalent MFSs, a representative MFS is chosen by selecting the PDB structure with the highest resolution. The present step is applied to metal-binding proteins only as CATH, SCOP and Pfam classifications are not available for nucleic acids. Hence, no equivalent or equistructural site is defined for nucleic acids.

2.2 Computational approaches to locate metal-binding proteins in proteomes

The -omics revolution faced bioinorganic chemistry with a new challenging perspective: the understanding of metalloproteomes, i.e. the entire set of metal-binding proteins encoded by organisms. The study of metalloproteomes can be approached at different levels of detail, spanning from the simple identification of metal-binding proteins to the more challenging comprehension of how metalloproteomes, together with all other cellular components, contribute to the metabolism of healthy cells and, under pathological conditions, lead to the onset of metal-associated diseases [48]. This latter level of knowledge builds upon

many intermediate studies, including the identification of metal sites and the definition of the native metal ions for all metal-binding proteins, as well as the structural/functional study of these systems. In the last decade, metalloproteomics has attracted the interest of an increasing number of scientists, who developed a portfolio of approaches to the investigation of metalloproteomes based on both experimental [49,50] and computational [50,51] methods.

Presently, experimentally available techniques have the general purpose of defining the complete set of metal-binding proteins encoded in genomes, and are largely based on modifications of classical proteomics and analytical tools. Without taking into account the specific limits of each technique, all current experimental approaches suffer from two main general limitations: (i) the native metal ion can be lost during protein manipulations (e.g. purification), especially in the case of transient binding sites, and (ii) non-native metal ions can bind in place of native ones, which may mislead the investigator also with respect to the function of the protein. On the other hand, computational approaches to metalloproteomics are generally designed to predict whether a sequence can bind a metal [52,53] and, in some cases, identify the metal site within the sequence [54-56]. These approaches are largely based on the development of models built on the 3D structural information available in the PDB and have exploited combined searches for known metal-binding domains and/or local sequence similarity to known metal-binding motifs [52,57] as well as supervised learning machines [51,54,56]. Computational approaches can complement experimental methods [58] by exploring wide amounts of sequences with very limited effort, in order to direct the more expensive experimental efforts. Consequently, various bioinformatics approaches have been developed to predict the metal-binding sites in a single sequence [59-61] but very few methods do allow metalloproteomics data analyses combined with the metal site prediction.

By exploting the information contained in MetalPDB, our group developed MetalPredator [62], a tool to predict iron–sulfur proteins from protein sequence, also at the whole proteome level. This tool integrates a domain-based approach with an approach designed to search for metal-binding motifs found in proteins with known structure. MetalPredator uniquely combines global and local searches to define whether a protein is a potential metal-binding protein. To validate the general methodology, the tool was firstly developed for the prediction of iron-sulfur clusters, showing good performances, both in terms of precision and recall.

3 AIM OF THE WORK

The general aim of my PhD project was to improve the knowledge about metal-binding proteins, focusing on the relationship between their structures and their sequences.

The primary method to identify a metal-binding protein is based on evidence derived from the presence of metals bound to the protein in the structure solved by experimental techniques. Therefore, in the first part of my project I worked on the upgrade of MetalPDB, which, as described above, is based on the structural information contained in the Protein Data Bank. In particular, I developed a protocol to identify apo sites (i.e., metal sites devoid of the metal cofactor) in protein structures, based on similarity to structurally characterized metal sites available in MetalPDB. Furthermore, I developed a new interface to some other tools to provide new features to MetalPDB, in order to increase the information available in it and to enhance the usability, usefulness and versatility of the resource by facilitating the access to the data. Also, I developed a completely new interface and a support database to manage the huge amount of data contained in MetalPDB database and the number of accesses to the web resource.

The protein sequences with associated structures are less than 1% of all known proteins. Therefore, in the second part of my project I worked on improving the prediction of metal-binding sites in protein sequences. In particular, I developed a new version of MetalPredator, making it able to predict iron-binding sites in proteins distinguishing among different iron cofactors, as well as zinc-binding and copper-binding sites. We used this tool to investigate the human portfolio of iron-proteins. Furthermore, I created a novel public resource called hMeProt, which collects data about human metal-binding proteins identified by predictive methods or experimental studies (metalloproteomics or structure determination). This new resource aims at integrating human metalloproteome data with other types of information so as to frame each metal-binding protein into the cellular/organismal context. From another perspective, the integration of data will produce a metal-centered view of the existing biological databases.

4 METHODS

4.1 MetalPDB version 2

The server side system for the new MetalPDB version (reported in section 4.1.1., 4.1.2, 4.1.3 and 4.1.4) were developed in Python; the front end was developed in HTML, Python and Javascript (by exploiting JQuery library). The framework used was Pylons.

For the functional annotation of MFSs the definition of the site classes are described in Table 4, while the functions associated to physiological sites are described in Table 5.

Table 4: *Descriptions of site classes*

Site Class	Description
Physiological Site	A site that has a confirmed physiological role. Each physiological site has an associated function (see Table 5).
Modified Physiological Site	At least one metal ion has been removed, added or substituted by another metal with respect to the physiological site. it may have more than one of this modifications: • A physiological metal ion is substituted by another one When the position of the native metal ion is filled by another metal without in vivo relevance. • A metal ion is removed When the physiological metal site has a vacant position (no metal ions occupy the position). • A metal ion is added When the physiological metal site presents a new position, occupied by an additional metal ion.

Not Physiological Site	A site for which current knowledge suggests that there is no physiological relevance within the cell. it may be: • Spurious The result of a binding event that is observed due to experimental procedures but is not relevant in vivo. • Artificial The result of a binding event occurring due to engineered or chemical modifications of the macromolecule. • Inhibitory The result of a binding event induced to inhibit the function of the protein for in vitro studies (the binding does not occur in vivo).
Unknown	The physiological relevance of the site is unknown. Unknown may be the: • Site When it is unknown if the site has a physiological relevance. • Metal occupancy When it is known that the site has a physiological relevance, but it is unknown which metal ion(s) occupies it in vivo.

Table 5: *Descriptions of functions associated to physiological sites (a function is associated to the site as a whole, not to each metal within the site)*

Function	Description
Catalytic	When the metal ion is directly involved in the reaction mechanism of the enzyme. • Redox When the metal ion partecipates to the reaction mechanism by donating/accepting electron(s). • Not Redox When the metal ion participates to the reaction mechanism, but maintains its redox number throughout the reaction.

Structural	When the metal ion stabilizes the 3D or higher-order structure of the macromolecule it may aim to stabilize: • The tertiary structure of a biomolecule • The quaternary structure of a protein • The complex interface of biomolecules
Transport	When the metal ion binds other chemical species that are then transported together with it and eventually released.
Electron Transfer	When the metal ion transports electrons.
Regulatory	When the metal ion is involved in controlling the activity of the system or in the regulation of cellular processes. it may control: • Catalysis When the binding of the metal ion enhances/inhibits the activity of an enzyme. • Expression When the binding of the metal ion induces/inhibits transcription.
Substrate	When the metal ion is the target of the protein. it may aim to: • Sense the presence of the metal or of a metal-containing cofactor • Transport the metal or a metal-containing cofactor • Store the metal or a metal-containing cofactor • Degradate a metal-containing cofactor • Biosynthesize a metal-containing cofactor
Protection	When the metal ion has the aim of preserving and defending a molecule from adverse reactions

4.1.1 Solvent accessibility and secondary structure information on the site

I added secondary structure and solvent accessibility to the precomputed analyses of the structural properties of MFSs. For each metal-binding protein, I used ProMotif [63]

(http://www.img.bio.uni-goettingen.de/ms-www/internal/manuals/promotif/promotif.html) to calculate the secondary structure elements of the entire 3D structure and then linked this information to the MFSs within the structure. The same procedure was applied with the program NACCESS (http://wolf.bms.umist.ac.uk/naccess/) to compute the solvent accessibility of the metal-binding residues in each MFS. For the calculation of solvent accessibility, each chain in the structure was considered individually and the steric hindrance of the metal neglected. The program provides the absolute and relative solvent accessibility for each residue; the relative values are calculated as the ratio between the absolute solvent accessibility value and that in an extended tripeptide (Ala-X-Ala) conformation.

4.1.2 FTP server and flat database

The FTP server allows the user to download all those sites that bind a specific metal ion. In this respect, a package of programs was developed to group together the MFSs and to move all the PDB files corresponding to these MFSs to the FTP server (each group is available as a compressed tar file). The FTP update was integrated in the updating process of MetalPDB. The flat database is available in the download section of MetalPDB and is provided in XML format.

4.1.3 Advanced search

I implemented an *Advanced search* page to query MetalPDB. This allows the user to submit a list of PDB codes and to choose the information about the MFSs of interest. After the submission of the PDB codes, the system checks them and reports all sites within the structures submitted. Then, the user can choose if he/she wants to select all the MFSs or select only some of them. Finally, the user can select the data of interest to result with a downloadable csv table, dynamically created, reporting all the information required. Examples of the information available about MFSs are: Pfam[47] domain, metal coordination geometry, donor atoms of the metal site.

4.1.4 Identification of potential metal-sites in apo-structures

The pipeline to identify apo-structures is composed by two main steps. The first step groups together all PDB chains having a sequence identity greater than 50% (cluster 50,

hereafter), by using the PDB's bc-50 file (ftp://resources.rcsb.org/sequence/clusters/bc-50.out). Cluster 50 include chains that do not contain the metal site and/or chains that, despite having the site, do not bind any metal ion (apo-structures). Then, all the chains within the same Cluster 50 are aligned and residues directly involved in the coordination of the metal ion(s) are mapped on the multiple alignment (Figure.9). The second step starts with the grouping of different Cluster 50 that contain equistructural MFSs, i.e. metal sites found in corresponding positions in similar structures (Equistructural groups, hereafter). All the alignments of Cluster 50 within the same Equistructural group where then aligned through T-Coffee [64] **.** When the number of sequences of an Equistructural group exceeded the maximum number of sequences managed by T-Coffee, only a subset of chains were kept from each Cluster 50. Finally, all the residues directly involved in metal coordination were mapped onto the final alignment to check if these residues are conserved also in those protein structures that do not bind metal ions within the site. Apo-structures that conserve the metal-binding pattern are likely able to bind a metal ion, so the apo-site was extracted from the pdb file and structurally aligned to the metal-containing MFSs of the group, though MetalS^2, a MetalPDB tool (http://metalweb.cerm.unifi.it/tools/metals2/) [65]. All the information derived during the pipeline was stored in the MetalPDB (Figure.10).

Figure 9. *First step of the pipeline to identify apo-structures. The pattern templates are objects composed by the metal-binding pattern and data about the structures associated to it.*

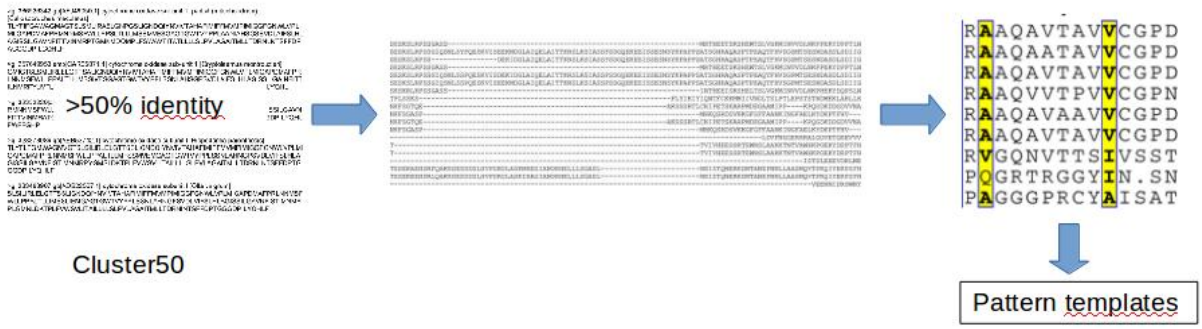
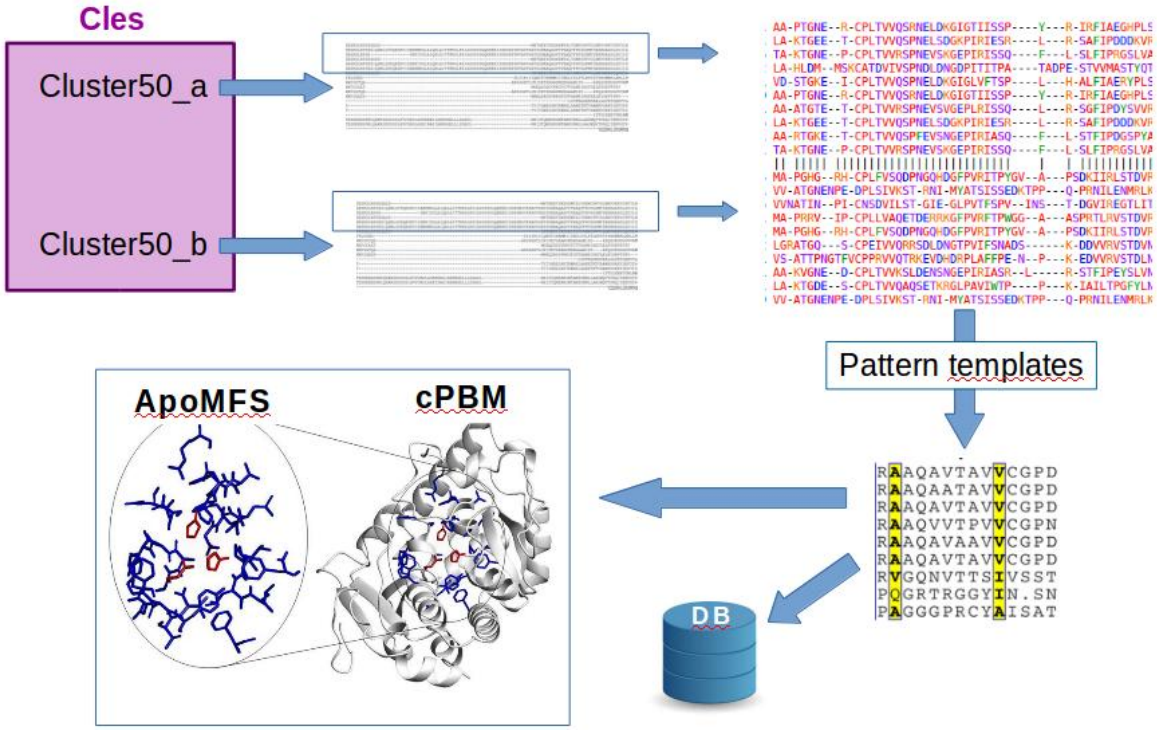


Figure 10. *Second step of the pipeline to identify apo-structures. Cles is an equistructural group, cPBM is an apo-structures that conserve the metal-binding pattern, apoMFS is an aposite.*



4.1.5 A NoSQL version of MetalPDB

To largely enhance the usability, usefulness and versatility of the resource by facilitating the access to the data, I implemented a new, NoSQL version of MetalPDB, using Mongo as database management system. The server side system for the new interface of MetalPDB was developed in Java language, the front end was developed in HTML, Scala, Javascript (by exploiting JQuery library) and the framework used was Play 2.4.

A package of Python programs extracts the core data from MetalPDB (version 2.0), then analyzes them, and finally inserts the processed data in the new NoSQL database. This guarantees an excellent database response performance, overcoming the limitations of the previous versions, ensuring rapid response times despite the increase of MetalPDB users. The new database will be automatically updated every time the MetalPDB (version 2.0) database is updated.

An example of a document from site collection of the new database is reported below.

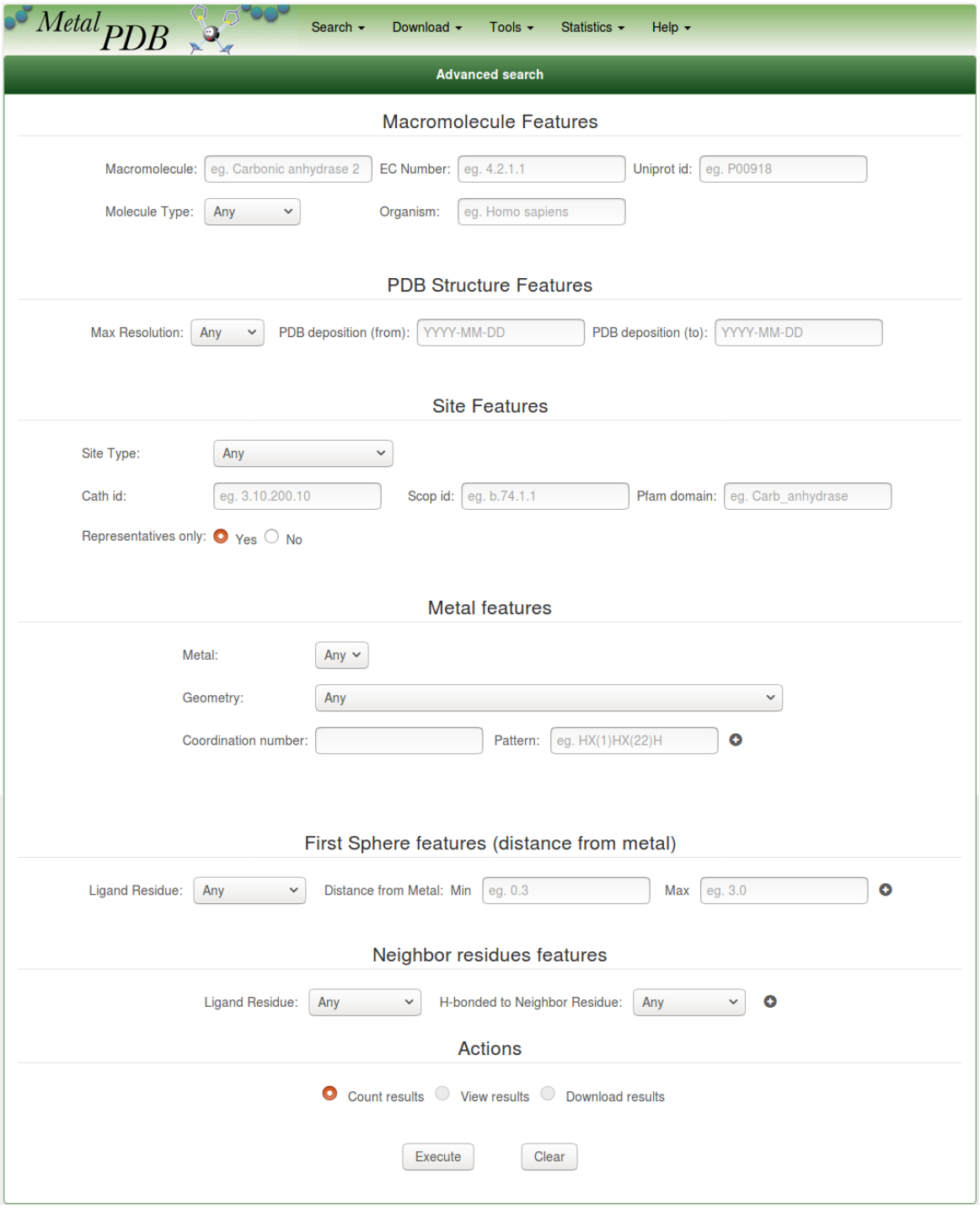

```
{
	"_id" : ObjectId("5d0b7c8ade671419baea4672"),
	"pfam" : [ "Carb_anhydrase" ],
	"pdb_code" : "12ca",
	"site_nuclearity" : 1,
	"name" : "12ca_2",
	"pdb_date" : "1991-10-01",
	"cath" : [ "3.10.200.10" ],
	 "scop" : [ "b.74.1.1" ],
	 "site_first_sphere_img" : "/first_sphere_images/12ca_2.png",
	"site_id" : "20",
	"scles_id" : 26865,
	"pdb_resolution" : 2.4,
	"location" : "Within a Chain",
	"fcles_id" : 35653,
	"site_image" : "/site_images/12ca_2.png",
	"is_representative" : false,
	"lig_num_str" : "HIS(3)",
	"site_type" : "Mononuclear",
	 "chains" : [{
			"molecule_type" : "protein",
			"chain_name" : "12ca_A",
			"molecule" : "Carbonic anhydrase 2",
			"ec_number" : "4.2.1.1",
			"uniprot" : "P00918",
			"organism" : "Homo sapiens"
		}],
	"metal" : [{
			"periodic_symbol" : "Zn",
			"metal_ligands_string" : "HIS_94(A),HIS_96(A),HIS_119(A)",
			"metal_id" : 26,
			"chain_letter" : "A",
			"ligand_res" : [{
					"donors" : [{
							"periodic_symbol" : "N",
							"atom_name" : "NE2",
							"side_main_chain" : "S",
							"atom_number" : 722,
							"distance" : 2.497244
						}],
					"solvent_accessibility_rel" : 15.6,
					"residue_name" : "HIS",
					"secondary_struct" : "L",
					"chain_letter" : "A",
					"endo_exo" : "endogenous",
					"residue_num" : 94,
					"solvent_accessibility_abs" : 28.53,
					"lig_id" : 755
				},
				{
					"donors" : [{
							"periodic_symbol" : "N",
							"atom_name" : "ND1",
							"side_main_chain" : "S",
							"atom_number" : 920,
							"distance" : 2.111593
						}],
					"solvent_accessibility_rel" : 2.2,
					"residue_name" : "HIS",
					"secondary_struct" : "E",
					"chain_letter" : "A",
					"endo_exo" : "endogenous",
					"residue_num" : 119,
					"solvent_accessibility_abs" : 4.09,
					"lig_id" : 756
```

```
                        },
                        {
                              "donors" : [{
                                          "periodic_symbol" : "N",
                                          "atom_name" : "NE2",
                                          "side_main_chain" : "S",
                                          "atom_number" : 743,
                                          "distance" : 2.112279
                                    }],
                              "solvent_accessibility_rel" : 2,
                              "residue_name" : "HIS",
                              "secondary_struct" : "S",
                              "chain_letter" : "A",
                              "endo_exo" : "endogenous",
                              "residue_num" : 96,
                              "solvent_accessibility_abs" : 3.74,
                              "lig_id" : 757
                        }],
                  "metal_info_string" : "ZN_262(A)_ZN",
                  "geometry" : "tetrahedron with a vacancy (regular)",
                  "pattern" : "HX(1)HX(22)H",
                  "res_number" : 262,
                  "exo_ligands" : "",
                  "atom_number" : 2029,
                  "res_name" : "ZN",
                  "coord_number" : 3,
                  "clem_position" : 1,
                  "periodic_name" : "Zinc",
                  "coord_code" : "tev",
                  "atom_name" : "ZN",
                  "first_sphere_img" : "/first_sphere_images/12ca_2_ZN_2029.png",
                  "endo_ligands" : "HIS_94(A), HIS_96(A), HIS_119(A)"
            }]
}
```


4.1.6 A new, more efficient, interface for MetalPDB

To allow quick access to data collected in MetalPDB, I implemented a new interface based on a Play framework application. The appearance of the web pages is similar to the first version of MetalPDB, as well as the major functionalities of the resource. The main difference with respect to the previous version is the development of a new advanced search (Figure. 11), which is organized in six sections: 1. macromolecule features, 2. PDB structure features, 3. site features, 4. metal features, 5. first sphere features, 6. neighbor residue features. Some of them allow to add more than one search block in the same section. Results of searches can be either visualized or downloaded as a custom report in the form of a csv file, and the user can select fields of interests to be included in the report.

Figure 11: *The logical operator used between the sections is "and", while between the blocks (if added) inside the same section is allowed the choice between "and" and "or".*



4.2 MetalPredator version 2.0

MetalPredator (http://metalweb.cerm.unifi.it/tools/metalpredator/) [62] is designed to predict metal-binding sites in protein sequence(s) at the whole proteome scale. The tool integrates an existing domain-based approach [66] with a new one designed to search for metal-binding motifs found in proteins with known structure, thus combining global and local searches to define whether a protein is a potential metal-binding protein.

To identify metal-binding sites in protein sequences, MetalPredator uses two libraries of Hidden Markov Model (HMM, hereafter) profiles that represent (1) Pfam[47] domains and (2) structural motifs binding metal ions. Metal-binding motifs are defined by splitting the Minimal Functional Sites (MFSs) stored in MetalPDB into fragments. Each fragment is a continuous stretch of protein sequence containing at least one metal ligand. The library of Pfam domains was built as described in [66]: it contains the profiles of both Pfam domains for which the metal ligands are known and domains annotated as metal binding but lacking information on the ligands. To build the library of motifs, each metal-binding sequence in MetalPDB was searched through PSI-Blast [67] into UniRef50 database [68]. All the hits with sequences in the output which conserved the metal ligands were then used to build a sequence profile of each fragment of the MFS contained in the input sequence.

MetalPredator uses the hmmscan tool [69] to match every input sequence to the profiles contained in the libraries. The predictions are based on the matching of the sequence with at least one profile and on the conservation of ligand residues on sequence (when they are known).

In its first version, MetalPredator was designed to predict iron-sulfur proteins; the pipeline to build libraries was time-consuming and each program was manually run; furthermore, the interface of the tool did not allow the user to perform flexible searches.

During my Ph.D. I developed a second version of MetalPredator, able to predict zinc-, copper- and iron-binding (including heme) sites. Since many Pfam domains are able to bind more than one metal within the same site, to refine the predictions based on the Pfam domains I developed a pipeline to build profiles of domains specific for each metal cofactor (for further details see par. 4.2.2). I also implemented an automatic pipeline which parallelizes the process of PSI-Blast[67] searches to reduce time required for the creation of HMM libraries (the jobs management was performed using PBS workload manager). Furthermore, I designed a new interface allowing users to select subset of libraries of sites; this interface dynamically creates

new HMM libraries based on user request. Finally, I worked at a stand-alone version of MetalPredator 2.0.

4.2.1 Creation of training datasets for iron- (heme and ions) zinc- and copper-proteins

In the first release of MetaPredator, aimed at predicting iron-sulfur sequences, libraries were built using a subset of all the iron-sulfur proteins. These were created by using as input for PSI-Blast[67] only iron-sulfur sequences having less than 50% of sequence similarity, in order to reduce time and the use of system resources. The subset was representative of whole population of iron-sulfur proteins because this class of metal-binding proteins does not show a large variability in ligand patterns. Instead, iron- (individual iron ions and heme), zinc- and copper-proteins show a much more ligand pattern variability so it is necessary to select at least one query protein for each different ligand pattern, even when they have a sequence identity greater than 50%. To this aim, I developed an algorithm composed of three main steps:

1. From MetalPDB, select all the protein chains that bind an input metal ion.
2. Cluster metal-binding chains with a sequence identity higher than 50% by using the PDB's bc-50 file (ftp://resources.rcsb.org/sequence/clusters/bc-50.out).
3. Perform a multiple sequence alignment of the protein chains within each cluster, using T-Coffee [64].
4. Map ligand residues on each sequence in the multiple alignment.
5. Select one sequence for each different pattern occurring in a cluster.

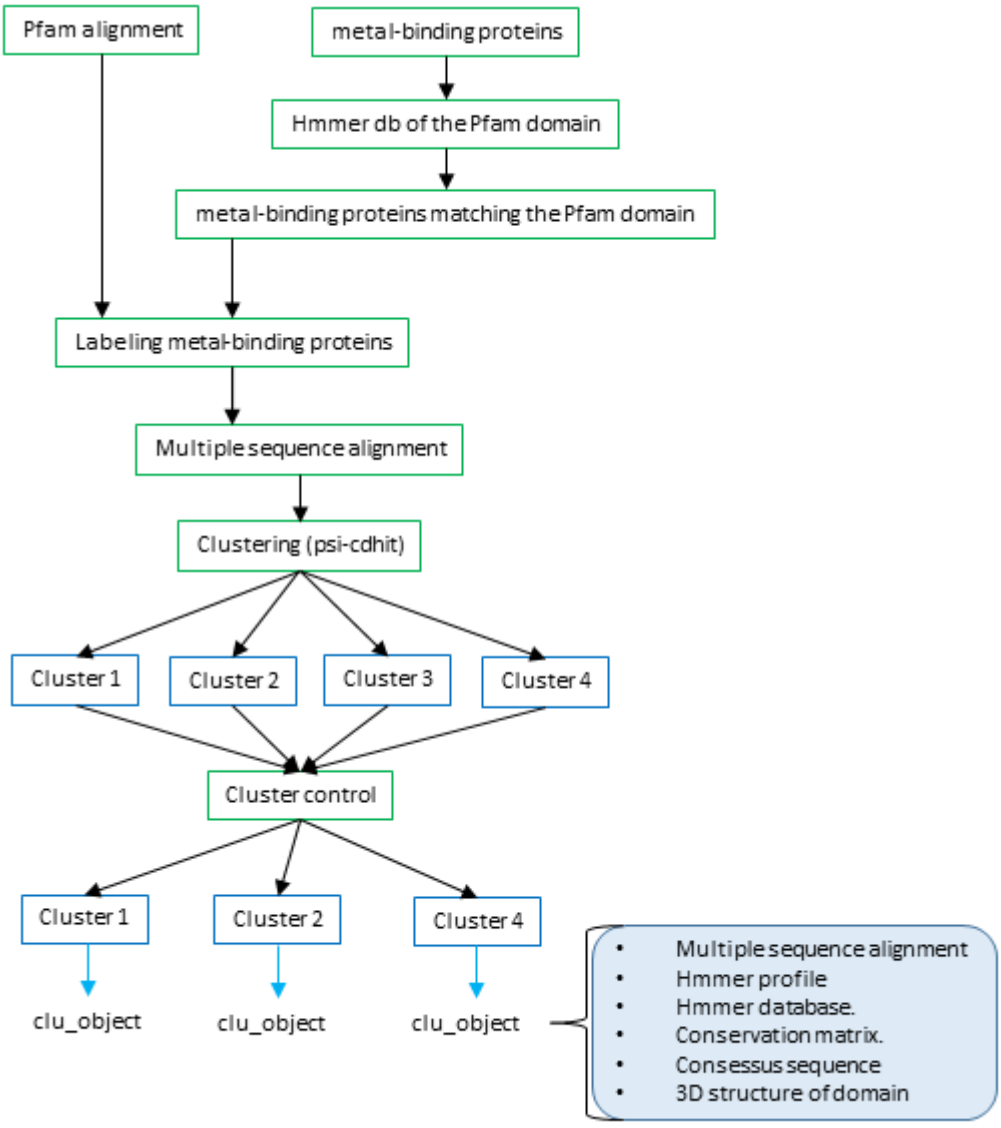
4.2.2 Development of a new pipeline to create specific profiles of Pfam-domains able to bind more than one metal within the same site

To refine the predictions of MetalPredator based on the Pfam[47] domains I developed a pipeline to build profiles of domains specific for each metal cofactor (Figure 12), based on clustering. The main steps are the follows:

1. Download of alignment of sequences on which the domain is built, as well as the Hmmer [70] profile of the domain and building of an Hmmer database.
2. Label each sequences in the alignment with the metal it contains.
3. Cluster aligned sequences in groups with 30% sequence identity, using psi-cdhit [71].

4. For each cluster containing at least 10 sequences:
    - Create a multiple sequence alignment
    - Build an Hmmer profile and create an Hmmer database.
    - Build the conservation matrix.
    - Write the consensus sequence (conservation rate > 0.7).
    - Build a 3D structure of the domain (from PDB structure if exists or from model) and color residues based on the conservation rate.

Figure 12. *Schema of pipeline to create specific profiles of Pfam-domains*



4.2.3 Test of the tool

To test the tool predictive performance I developed two protocols: one for the sensitivity and one for the specificity. The protocol for sensitivity, using Blast [72], clusters all the sequences that bound the same metal with at least 25% of identity. Then it tests the prediction on one sequence for each cluster using the sequences of all the other clusters to build the libraries. The protocol for specificity use a dataset composed by proteins known binding one of the follows metals: Zn, Mg, Co, Ca, Na, Cu, Fe-S, Mo, Ni, Mn. It clusters the metal-binding sequences with at least 25% of identity (using Blast [72]) and for each metal cofactor tests the prediction on one sequence for each cluster, excluding the clusters which include proteins binding cofactor on test. So, this protocol reports also which metals can be more easily confused by the MetalPredator prediction.

4.3 hMeProt

4.3.1 Methods to identify the metal-binding proteins

We used the human proteome provided by UniProt [44] (one protein sequence for each gene) to describe the human metalloproteome, i.e. the entire set of metal-binding proteins encoded by humans. In hMeProt, metal-binding capabilities are identified through the application of five methodologies that, however, are different both in the reliability of the annotation/prediction and in the level of details for the metal-binding protein (two of them simply identify metal-binding proteins while the remaining three methodologies are able to identify also the residues that directly coordinate the metal ion). For the above reasons, the protocol provides a hierarchy of methods (from the most reliable and detailed to the least reliable).

- **Manual annotation of the metal site in the MetalPDB entry**: The site is identified using data taken from the MetalPDB database, i.e. from the protein structure. When the site is annotated as a "physiological metal site" by MetalPDB curators, then the structure is used to identify the residues directly involved in the metal coordination. The MFS is directly retrieved from MetalPDB.
- **Manual annotation of the metal site in the Uniprot entry**: The metal site is identified using the "sequence features" section of the relative UniProtKB entry, which contains manual curated annotations that describe the residues that directly coordinate the metal ion of interest. When the protein structure is not available, if possible, the software calculates a 3D model of the protein to extract the putative MFS (without the metal bound).
- **Prosite method**: The site is identified using a protocol which integrates the Prosite pattern with the structural information contained in MetalPDB. This method uses libraries of pattern profiles specific for each metal type, built by the following protocol:
    1. Scan all known physiological metal-binding structures in MetalPDB with the PS_SCAN software to find Prosite patterns within their associated sequences [73].
    2. Select those patterns which include and conserve the metal-binding residues.
    3. Map metal-binding residues position on the Prosite pattern.

4. This method is able to predict the metal-site within the protein sequence. When the protein structure is not available, if possible, the software calculates a 3D model of the protein to extract the putative MFS (without the metal bound).

- **Uniprot no ligands method**: The metal-binding protein is identified using the "cofactor" section of the relative UniProtKB entry. This annotation, based on the literature, just provides the type of the metal bound to the protein. No information is available for the metal-site.
- **Gene Ontology method**: The metal-binding protein is identified on the basis of the Gene Ontology [74] annotation. This annotation, based on the literature, just provides the type of the metal bound by the protein. No information is available for the metal-site.

Each UniProt sequence was aligned to the corresponding structure by using Nwalign (http://zhanglab.ccmb.med.umich.edu/NW-align), Each method works independently, so one site can be predicted by more of one method.

4.3.2 hMeProt database

We applied the methodologies described above to predict the human metalloproteome. hMeProt was designed using a non-relational approach (noSQL), in order to have quick access to the information from the interface of the web resource. As database management system was used MongoDb. The MeProt database was designed to optimize the management, update and access to data. In this respect, I developed software tools to automatically maintain the data up-to-date. The MeProt database was integrated with various other biological resources to associate each metal-binding protein with the largest possible amount of information available, with the aim of facilitating the process of knowledge discovery by the users. Each metal-binding protein is identified by the UniProt[44] identifier and is associated with various types of data such as cellular localization, metabolic pathways, and genetic variations. Biological resources used to integrate data in hMeProt are reported below:

- UniProtKB was used to get general information on each human protein.
- The Human Protein Atlas [75] was used to get data on the expression of genes in tissues, on the expression levels in the cellular type of each tissue, on protein subcellular locations, on genes used as prognostic marker for cancers.
- SwissVar [76] was used to get data about variants on protein sequences.

- dbSNP [77] was used to get data about variants (only single-nucleotide polymorphism) on protein sequences.
- ClinVar [78] was used to get data about the clinical significance of single-nucleotide polymorphisms on protein sequences.
- KEGG Pathway database and KEGG BRITE database [79] were used to get data on metabolic pathways.
- NCBI Gene [80], OMIM [81], MedGen [82] and Orphanet [83] were used to get data on pathologies.
- Gene Ontology [74] annotation was used to get data on protein function.
- PDB was used to get data on the available protein structures.
- Protein Model Portal [84] and SWISS-MODEL [85] were used to get data on the available 3D models of proteins.

Every time hMeProt is updated, all the above resources are newly queried to obtain updated data.

An example of a document from protein collection is reported below (the symbol {…} defines subdocuments of arrays with more than three items).


```
{
	"_id" : ObjectId("5bd06dd7a17ddd07e39e37b0"),
	"uniprot_secondary_ac" : [ "B2R7G8","Q6FI12","Q96ET9" ],
	"sequence" :
"MSHHWGYGKHNGPEHWHKDFPIAKGERQSPVDIDTHTAKYDPSLKPLSVSYDQATSLRILNNGHAFNVEFDDS

QDKAVLKGGPLDGTYRLIQFHFHWGSLDGQGSEHTVDKKKYAAELHLVHWNTKYGDFGKAVQQPDGLAVLGIF

LKVGSAKPGLQKVVDVLDSIKTKGKSADFTNFDPRGLLPESLDYWTYPGSLTTPPLLECVTWIVLKEPISVSS
                    EQVLKFRKLNFNGEGEPEELMVDNWRPAQPLKNRQIKASFK",
	"taxonomy" : " Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Euteleostomi;
                     Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
Catarrhini;
                     Hominidae; Homo.",
	"sub_location" : [ "Cytoplasm","Cell membrane" ],
	 "metals" : [ "Zn" ],
	"reviewed" : true,
	"ec_numbers" : [ "4.2.1.1" ],
	"gene_synonyms" : [ ],
	"uniprot_ac" : "P00918",
	"sites" : [{
			"cluster_id" : "1",
			"methods" : [ "uniprot_features","metalpdb" ],
			"metals" : "Zn",
			"cofactors" : ""
		},
		{
			"cluster_id" : "0",
			"methods" : "uniprot_cofactor",
```

```
                        "metals" : "Zn"
                },
                {
                        "cluster_id" : "0",
                        "methods" : "go",
                        "metals" : "Zn"
                }
        ],
        "cross_references" : [{
                        "source" : "Pfam",
                        "ids" : [ "PF00194" ]
                },
                {
                        "source" : "RefSeq",
                        "ids" : [ "NP_000058.1" ]
                },
                {
                        "source" : "Reactome",
                        "ids" : [ "R-HSA-1237044","R-HSA-1247673","R-HSA-1475029"]
                }, {…}],
        "function" : "Essential for bone resorption and osteoclast differentiation
(By similarity). Reversible hydration of carbon differentiation (By similarity).
Reversible hydration of carbon dioxide. Can hydrate cyanamide to urea. Involved in
the regulation of fluid secretion into the anterior chamber of the eye. Contributes
to intracellular pH regulation in the duodenal upper villous epithelium during
proton-coupled peptide absorption. Stimulates the chloride-bicarbonate exchange
activity of SLC26A6.",
        "recommended_names" : [ "Carbonic anhydrase 2" ],
        "variants" : [{
                        "pathologic" : true,
                        "description" : "in OPTB3; in Czechoslovakia.",
                        "FTId" : "VAR_001381",
                        "position" : "92",
                        "pathologies" : "osteopetrosis, autosomal recessive 3",
                        "aa_2" : "P",
                        "aa_1" : "Q"
                },
                {
                        "pathologic" : true,
                        "description" : "in OPTB3; partial loss of activity.",
                        "FTId" : "VAR_021009",
                        "position" : "94",
                        "pathologies" : "osteopetrosis, autosomal recessive 3",
                        "aa_2" : "Y",
                        "aa_1" : "H"
                },{…}],
        "SNPs" : [{
                        "description" : "in Jogjakarta;",
                        "OMIM" : null,
                        "clinical_significance" : "Pathogenic",
                        "position" : "18",
                        "dbSNP" : "rs118203931",
                        "pathologies" : "CARBONIC ANHYDRASE II VARIANT",
                        "aa_2" : "E",
                        "aa_1" : "K",
                        "MedGen" : null
                },
                {
                        "description" : "in OPTB3; frequent mutation;",
                        "OMIM" : "259730",
                        "clinical_significance" : "Pathogenic",
                        "position" : "107",
                        "dbSNP" : "rs118203933",
                        "pathologies" : "Osteopetrosis with renal tubular acidosis",
                        "aa_2" : "Y",
                        "aa_1" : "H",
                        "MedGen" : "C0345407"
                }, {…}],
        "length" : 260,
```

```
      "sub_note" : " Colocalized with SLC26A6 at the surface of the cell membrane
in order to form a bicarbonate transport metabolon. Displaced from the cytosolic
surface of the cell membrane by PKC in phorbol myristate acetate (PMA)-induced
cells.",
      "organism" : "Homo sapiens (Human).",
      "alternative_names" : [ "Carbonate dehydratase II","Carbonic anhydrase
C","Carbonic
                              anhydrase II" ],
      "gene_name" : "CA2",
      "variants_in_first" : [ "Zn" ],
      "variants_in_second" : [ "Zn" ],
      "variants_count" : 7,
      "cell_compartments" : [ "Cytoplasm","Cell membrane" ],
      "atlas_tissue" : [{
                  "cell_type" : "glandular cells",
                  "tissue" : "appendix",
                  "reliability" : "Enhanced",
                  "level" : "High"
            },
            {
                  "cell_type" : "hematopoietic cells",
                  "tissue" : "bone marrow",
                  "reliability" : "Enhanced",
                  "level" : "Medium"
            }, {…}],
      "diseases" : [{
                  "prognostic_marker" : "favourable",
                  "p_value" : 0.00000557,
                  "pathology" : "renal cancer",
                  "sources" : [ "HPA_CA2" ]
                   },
            {
                  "sources" : [ "MedGen_C0345407","OMIM_259730" ],
                  "pathology" : "Osteopetrosis with renal tubular acidosis"
            }],
      "drugs" : [{
                  "drug_name" : "Acetazolamide",
                  "kegg_drug_id" : "DG01134"
            },
            {
                  "drug_name" : "Brinzolamide",
                  "kegg_drug_id" : "D00652"
            }, {…}],
      "pathways" : [{
                  "kegg_orthology_2" : "Nitrogen metabolism",
                  "kegg_orthology_1" : "Energy metabolism",
                  "kegg_orthology_0" : "Metabolism",
                  "kegg_pathway_id" : "hsa00910",
                  "pathologic_pathway" : false,
                  "pathway" : "Nitrogen metabolism"
            },
            {
                  "kegg_orthology_2" : "Proximal tubule bicarbonate reclamation",
                  "kegg_orthology_1" : "Excretory system",
                  "kegg_orthology_0" : "Organismal Systems",
                  "kegg_pathway_id" : "hsa04964",
                  "pathologic_pathway" : false,
                  "pathway" : "Proximal tubule bicarbonate reclamation"
            }, {…}]
}
```


An example of a document from metal_site collection is reported below (the symbol {…} defines subdocuments of arrays with more than three items).

```
{
	"_id" : ObjectId("5c7686a7de67140a7341a812"),
	"uniprot_ac" : "P00918",
	"metal" : {
		"note" : "Zinc",
		"symbol" : "Zn"
	},
	"evidence" : "Experimental evidence",
	"pdb" : [{
			"code" : "1FQN",
			"resolution" : 2,
			"interval" : "1-260"
		},
		{
			"code" : "1FQL",
			"resolution" : 2,
			"interval" : "1-260"
		},
		{
			"code" : "1FQM",
			"resolution" : 2,
			"interval" : "1-260"
		},
		{
			"code" : "1FQR",
			"resolution" : 2,
			"interval" : "1-260"
		},
		{
			"code" : "1BIC",
			"resolution" : 1.9,
			"interval" : "2-260"
		}],
	"first_sphere" : [ 94,96,119 ],
	"method" : "uniprot_features",
	"metal_pattern" : "H_94,H_96,H_119",
	"cluster_id" : "1",
	"second_sphere" : [
7,62,65,66,67,92,93,95,97,98,104,105,106,107,115,116,117,118,120,
                          121,142,143,144,198,199,208,243,244 ],
	"ligands_pattern" : "HXHX(22)H"
}
```


4.3.3 Web resource technical overview

The web application back-end was developed in Java language, Play 2.6 was used as framework for the web application interface, which was developed in Scala, javascript and HTML. The charts are dynamically designed using GoogleChart API, that allows to create interactive graphs.

5 RESULTS

5.1 MetalPDB

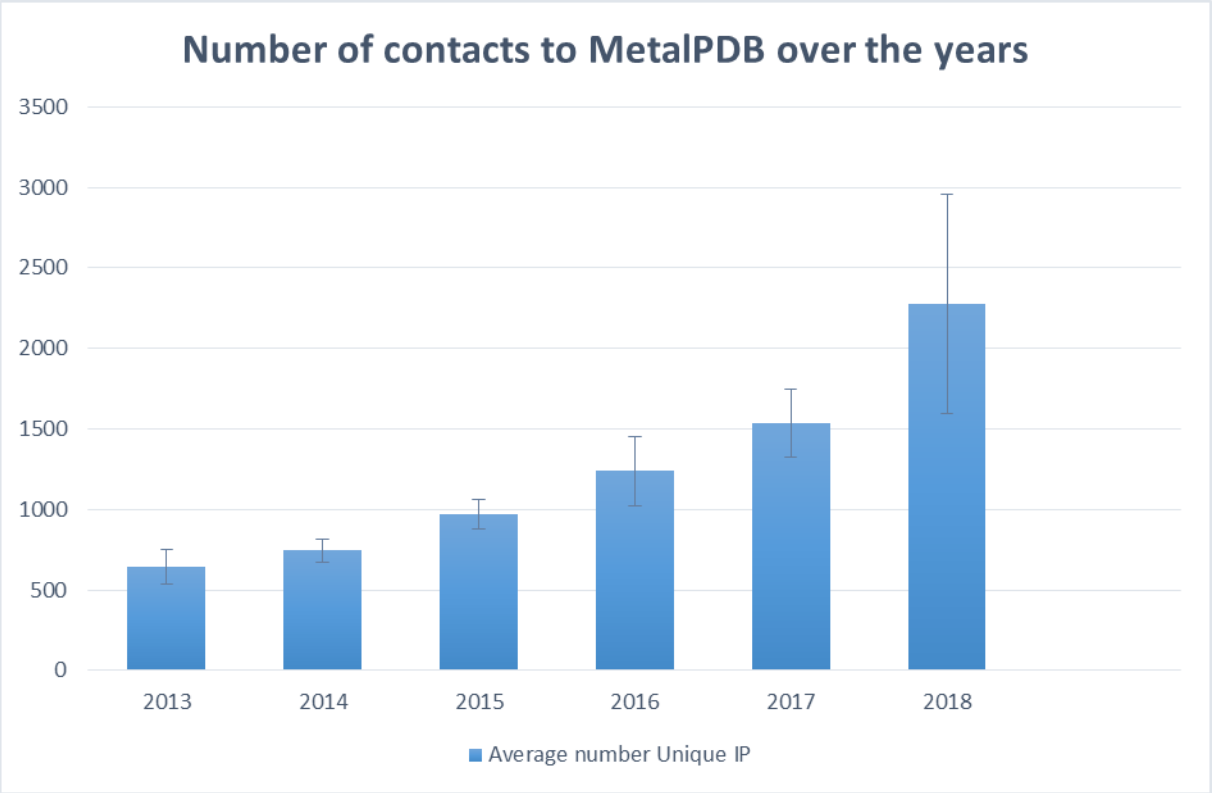
5.1.1 MetalPDB in 2018

MetalPDB (http://metalweb.cerm.unifi.it) is an important resource in the field of bioinorganic chemistry, as the number of online tools and databases dedicated to metals in biology is scarce with respect to the size of the scientific challenge. The database collects and allows easy access to the knowledge on metal-binding proteins, exploiting the structural information on metal sites stored in the PDB. In MetalPDB, metal sites are stored as Minimal Functional Site (MFS) objects, i.e. the local structure of the site including the metal ion or cofactor, its ligands and any other atom belonging to a chemical species within 5 Å from a ligand (Figure 4). By construction, MFSs contain the bulk information on the factors tuning the affinity of a site for its native metal versus other ions. Similarly, MFSs include the structural factors determining the chemico-physical, and consequently the functional, properties of the metal ion. MetalPDB groups MFSs in clusters of equivalent sites, i.e. sites in which the same metal cofactor is located in the same position of proteins sharing the same structure.

The architecture of the database is based on an accurate structural classification of metal sites and of metal-binding proteins containing them, allowing users to perform flexible and detailed queries and analyses, and facilitating its management and update. This resource provided the scientific community with an unprecedented picture of the entire landscape of known metal-binding proteins, also thanks to the statistic section included in the web-interface. The thoroughness of MetalPDB makes it useful for large-scale studies on interaction of metals with biological macromolecules, for example at the level of whole proteomes.

For the above reasons, it is not surprising that since its first publication (seven years ago), MetalPDB has met an ever-increasing interest from the scientific community (Figure 13). In its second release (on which I worked during my Ph.D.) the resource has reached an average of 4000 visits per month (a new visit is counted if the same IP makes requests at half-hour intervals or longer). It is kept constantly updated and at present it contains 297.153 sites from 53.366 structures.

Figure 13. *The growth of the unique IPs contacting the MetalPDB database.*



Published online 25 October 2017 *Nucleic Acids Research, 2018, Vol. 46, Database issue* ***D459–D464***
doi: 10.1093/nar/gkx989

MetalPDB in 2018: a database of metal sites in biological macromolecular structures

Valeria Putignano[1], Antonio Rosato[1,2], Lucia Banci[1,2] and Claudia Andreini[1,2,*]

[1]Magnetic Resonance Center (CERM)—University of Florence, Via L. Sacconi 6, 50019 Sesto Fiorentino, Italy and [2]Department of Chemistry—University of Florence, Via della Lastruccia 3, 50019 Sesto Fiorentino, Italy

Received September 14, 2017; Revised October 10, 2017; Editorial Decision October 11, 2017; Accepted October 20, 2017

ABSTRACT

MetalPDB (http://metalweb.cerm.unifi.it/) is a database providing information on metal-binding sites detected in the three-dimensional (3D) structures of biological macromolecules. MetalPDB represents such sites as 3D templates, called Minimal Functional Sites (MFSs), which describe the local environment around the metal(s) independently of the larger context of the macromolecular structure. The 2018 update of MetalPDB includes new contents and tools. A major extension is the inclusion of proteins whose structures do not contain metal ions although their sequences potentially contain a known MFS. In addition, MetalPDB now provides extensive statistical analyses addressing several aspects of general metal usage within the PDB, across protein families and in catalysis. Users can also query MetalPDB to extract statistical information on structural aspects associated with individual metals, such as preferred coordination geometries or aminoacidic environment. A further major improvement is the functional annotation of MFSs; the annotation is manually performed via a password-protected annotator interface. At present, ∼50% of all MFSs have such a functional annotation. Other noteworthy improvements are bulk query functionality, through the upload of a list of PDB identifiers, and ftp access to MetalPDB contents, allowing users to carry out in-depth analyses on their own computational infrastructure.

INTRODUCTION

For the large majority of organisms, 30–40% of proteins require one or more metal ions to perform their biological function in cells (1;2). Additionally, metal ions play a decisive role in stabilizing the structure of nucleic acids (3). MetalPDB (4) is a resource derived from the automated analysis of all the three-dimensional (3D) structures of the adducts between biological macromolecules and metal ions or metal-containing cofactors available from the Protein Data Bank (PDB, http://www.wwpdb.org/) (5). MetalPDB stores the metal sites observed in PDB structures in the form of Minimal Functional Sites (MFSs) (6;7). Each MFS is the ensemble of atoms of the metal cofactor, the metal ligands and any other residue or chemical species within 5 Å from a ligand. The MFS describes the local 3D environment around the cofactor, independently of the larger context of the macromolecular structure in which it is embedded. The usefulness of the MFS concept has its chemicophysical foundation in the fact that the local environment of the metal has a determinant role in tuning its properties and thus its chemical reactivity. Consequently, MFSs can provide an unbiased insight into the function or mechanism of action of a metalloprotein (i.e. a protein that binds at least one metal ion or metal-containing cofactor) (6;8). The structural comparison of MFSs is useful also to predict function from 3D structure in the absence of experimental biochemical data. MetalS^3 tool is designed to search MetalPDB for all those sites that have a similar local structure with a query site (9).

Since its first release, in 2012, MetalPDB has been widely exploited by the scientific community. In the last 12 months, there have been on average 1450 unique IPs contacting the database each month, corresponding on average to almost 4000 visits (a new visit is counted if the same IP makes requests at half-hour intervals or longer). The current release includes 287 122 sites from 50 797 structures. It was 175 115 in the first release of MetalPDB (64% growth in 6 years). MetalPDB is updated monthly in an automated manner.

In the current update of MetalPDB, we extended its contents to include various new features and expanded the information available via the web interface. A number of improvements were made to the usability of the web interface, including bulk query functionality and faster visualization of pages. As a major upgrade, we specifically addressed the identification of potential MFSs in 3D structures lacking the metal cofactor. In addition, statistical analyses on the MetalPDB contents are now available on the web site, in order to provide a better understanding of the diversity of the biochemical roles of metals.

*To whom correspondence should be addressed. Tel: +39 55 4574267; Email: andreini@cerm.unifi.it

New contents of MetalPDB

We added secondary structure and solvent accessibility to the precomputed analyses of the structural properties of MFSs. For each metalloprotein, we used ProMotif (http://www.img.bio.uni-goettingen.de/ms-www/internal/manuals/promotif/promotif.html) (10) to calculate the secondary structure elements of the entire 3D structure and then linked this information to the MFSs within the structure. The same procedure was applied with the program NACCESS (http://wolf.bms.umist.ac.uk/naccess/) to compute the solvent accessibility of the metal-binding residues in each MFS. For the calculation of solvent accessibility, each chain in the structure was considered individually and the steric hindrance of the metal neglected.

We introduced functional annotations for MFSs. All equivalent sites (i.e. MFSs that occur at the same position within a conserved protein fold, as observed in the structural alignment of all the chains of the superfamily, and bind the same metal ions) share the same functional annotation so the clustering procedure is critical for the quality of annotation. To improve the homogeneity of groups of equivalent sites we revised our previous procedure (4) (see point 7 of the Section Database Construction) by using exclusively the Pfam domain classification (11) as the criterion to create protein superfamilies. Functional annotations are manually curated via a dedicated, password-protected annotator interface. This interface uses drop-down menus and a guided annotation procedure in order to minimize clerical errors. At the top level, we annotate the physiological relevance of each MFS by assigning it to one of these classes: 'Physiological site', 'Modified Physiological site', 'Not physiological site' and 'Unknown' (for a description, see http://metalweb.cerm.unifi.it/help/functional_annotation/). A MFS is considered to be physiological only if all the metal ions identified in the structure correspond to those required for the system to function in the cell (native metal ions), and all and only the required metals are present. In a modified physiological site, at least one metal ion has been removed, added or substituted by another metal with respect to the physiological site. A not physiological site is one that is known to be not relevant *in vivo*. When a metal ion in a structure has no donor atoms in its first coordination sphere it is automatically annotated as 'Not physiological'; this can happen e.g. if a water molecule in the crystal structure was incorrectly assigned as a metal by the depositors. Each physiological site has one or more associated functions among 'Catalytic', 'Structural', 'Transport', 'Electron transfer', 'Regulatory', 'Substrate' and 'Protection' (see http://metalweb.cerm.unifi.it/help/functional_annotation/) (12). Some of these terms have a further level of annotation to improve the information content of the record. At present, a functional information is available for the majority of the sites binding iron or copper (Table 1).

A commonly asked question is what the structural impact of metal-binding is at the local and/or global structural level. To address this the 3D structures of the same protein with and without the cofactor needs to be compared. We therefore implemented a protocol to identify protein structures related to a structurally characterized MFS available in MetalPDB but devoid of the metal cofactor (apo-structures). To this end, we generated a multiple sequence alignment between all chains that bind equistructural MFSs (i.e. MFSs that occur at the same position within a conserved protein fold, regardless of the chemical identity of the bound metal) and the chains of apo-structures that have at least 50% identity with at least one of them. Potential MFSs in apo-structures are then identified based on the conservation of all metal-binding residues in this alignment. This procedure identifies apo-structures with the metal-binding pattern (Figure 1). Chains lacking one or more of the metal-binding residues probably have lost or significantly changed their interaction with the metal cofactor, and are listed separately as apo-structures without the metal-binding pattern (Figure 1). This provides the user with an innovative structural perspective on apo-structures, enabling the systematic analysis of the structural impact of metal binding and providing hints on the possible evolution of the MFS itself. In implementing this protocol, we realized that distinct groups of equistructural sites sometimes have some or even all metal ligands in common in the protein sequence alignment. Different groups of equistructural MFSs are created when structures with the same metal-binding protein domain have MFSs in different relative positions within the structural alignment of all the chains (4). However, the present sequence alignments reveal that this can happen while maintaining some metal ligands from the protein unchanged, i.e. the spatial shift of the MFS can be a result of structural rearrangements or flexibility rather than of evolutionary changes altering the sequence. We thus decided to dub sites that belong to different equistructural groups but share at least a protein ligand in the sequence alignment as 'related sites'.

Table 1. Percentage of annotated MFSs, grouped by metal. Data are shown only for essential metals (18)

Metal ion	Percentage of annotated sites
Cu	90%
Fe	86%
Mg	70%
Ni	35%
Mn	34%
K	32%
Na	29%
Mo	22%
Co	21%
Zn	17%
W	12%
Ca	12%
V	3%

This percentage reports on the number of MFSs with a functional annotation of any type with respect to the total number of MFSs in MetalPDB.

The MetalPDB interface

The 2018 version of MetalPDB features an additional mode of querying the database, i.e. by providing a list of PDB identifiers. The interface analyses the list to separate entries corresponding to metal-containing, apo- or not-metal-binding structures, and then allows the user to select specific MFSs from each metal-containing entry. In this way one

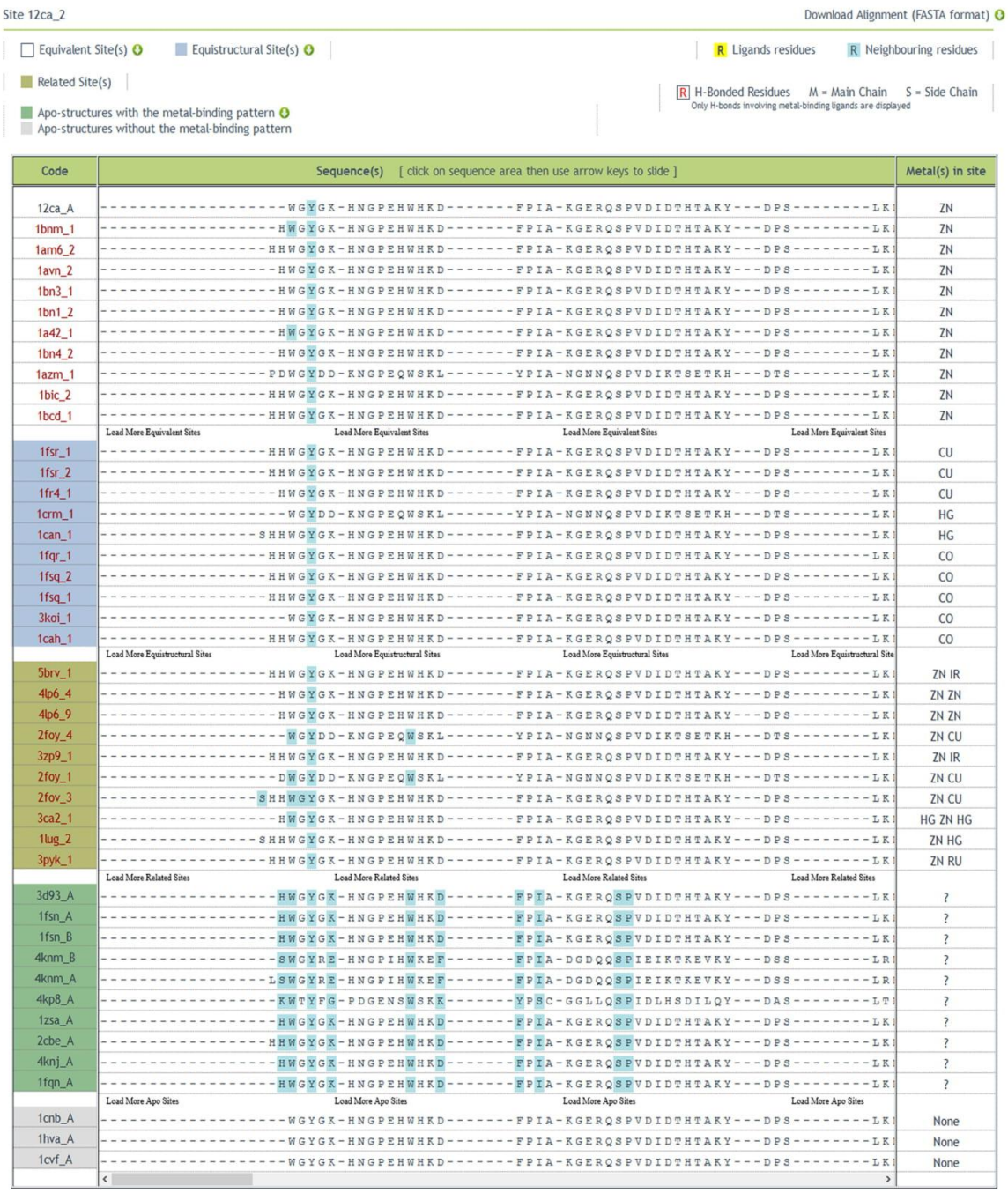


Figure 1. The Sequence tab for entry 12ca_2 (16). The new Sequence tab displays the sequence alignment of all proteins in the same superfamily. The proteins are grouped based on the relationship of their MFSs to the query MFS (equivalent or equistructural sites), whereas for proteins lacking any metal in the site the grouping is based on the conservation of the metal ligands (apo-structures with or without the metal-binding pattern). The metal ligands have a yellow background; the residues belonging to the MFS have a cyan background. H-bonded residues are highlighted in red.

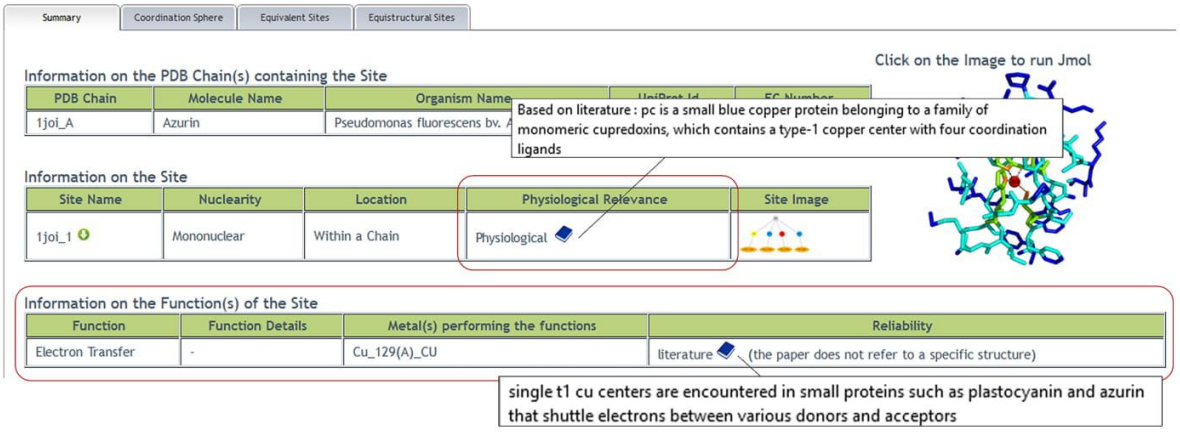


Figure 2. The summary page of 1joi_1 (17). The new *Information on the site* table reports the *Physiological relevance* of the site (highlighted with a red circle). Each physiological site has an associated function that is detailed in a further new table (*Information on the function(s) of the site*, also highlighted with a red circle) When an annotation is based on the literature, it is possible to display the sentence of the article that supports the functional annotation by hovering the mouse on the book icon. The book icon links to the article entry on PubMed.

can select, for example, only physiologically relevant sites or only a given site in a family of metalloproteins containing multiple MFSs. After completing the selection, it is possible to create a personalized report on the properties of all the selected MFSs. For each MFS, the report can include features of the site (CATH (13)/SCOP (14)/Pfam (11)) domain containing the site, number of ligands, EC number for metalloenzymes), of the metal (coordination geometry, coordination number, metal-binding pattern) and of the ligands (donor atoms, metal-to-donor distances). The report can be downloaded as a csv file.

To facilitate the analysis of the entire MetalPDB contents, we implemented two new options for large data download: an ftp interface providing access to all the MFSs, grouped by the bound metal (each group is available as a compressed tar file), and a link to a flat file version of the database.

MetalPDB returns results on a per-MFS basis, i.e. the result page shows the information contained in the database for an individual MFS. The information is distributed under different tabs within the page. Below we report the modified or the new tabs of the current version of MetalPDB:

- *Summary tab*: the table 'Information on the Site' now reports, when available, the physiological relevance of the site. When a site is 'Physiological', it also has an associated function, which is reported in the 'Function Details' table below. By hovering the mouse over the book icons, a sentence of the article supporting the annotation appears in a box (Figure 2). For Modified Physiological MFSs, we additionally provide a description of the changes with respect to the physiological site in a separate 'Site Modification' tab (see below).
- *Coordination Sphere tab*: each ligand is now associated with a relative solvent accessibility and with a secondary structure element.
- *Sequence tab*: this tab was not present in the previous version. It displays the sequence alignment of all the members of the protein superfamily of the query MFS (Figure 1). These include: (i) sequences harbouring equivalent sites (white) and (ii) sequences harbouring equistructural sites (blue); (iii) sequences with 'related sites' (light green), (iv) sequences of apo-structures that conserve all the metal-binding residues of the query MFS (dark green) and (v) sequences of apo-structures which have lost at least one metal-binding ligand with respect to the query MFS (grey). A structural superposition of the putative sites in the apo-structures with the metal-binding pattern to the query MFS can be downloaded. It is also possible to download the alignment of all the sequences. The user can move along the alignment by shifting it right or left to inspect its different regions, or by showing more or less sequences. A color code highlights the protein residues forming the MFS as well as the position and interactions of the metal-binding residues.
- *Site Modification tab:* this new tab is present when the query MFS is annotated as a 'Modified Physiological Site'. It details the modifications of the query MFS with respect to the physiologically relevant site(s).

Statistics pages

We have extended the previous version of MetalPDB to provide extensive statistics on its contents, providing both structural and functional information. Several different pages, which can be accessed via the *Statistics* drop-down menu of the navigation bar, are available:

- *Summary*, which lists the number of sites, atoms and PDB structures contained in MetalPDB on a per-metal basis;
- *Metals in PDB*, which provides an overview of the fractional occurrence of metal-binding structures in different repositories or for different macromolecule types, a histogram of the number of MFS with a given nuclearity (number of metal ions per site), and a statistics of the most common coordination geometries observed in MetalPDB;

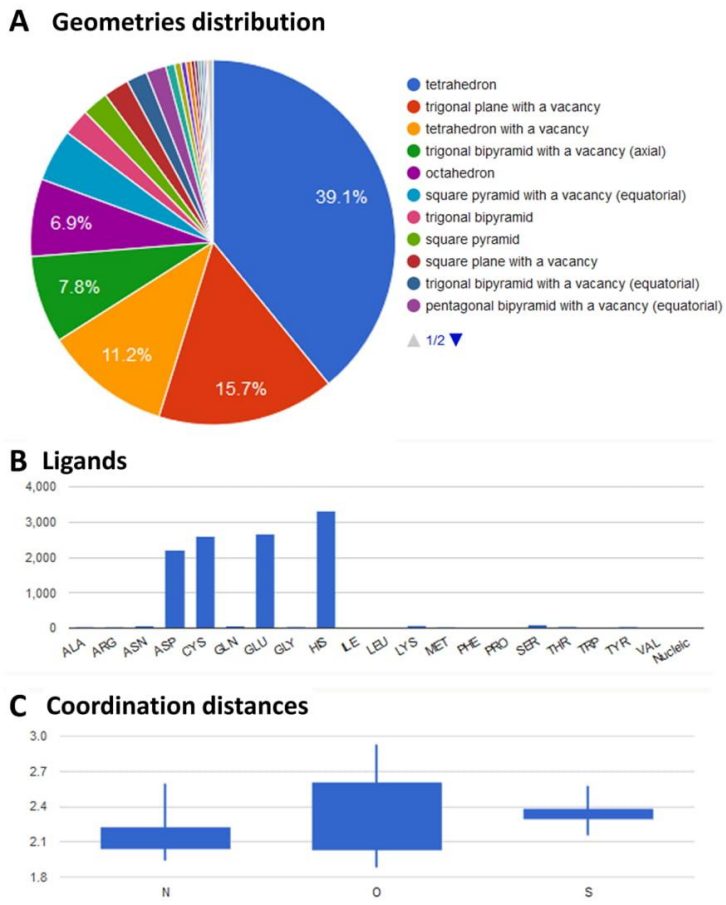


Figure 3. Example of statistics for zinc coordination spheres in the PDB. The information is accessible from the 'Per metal' statistics menu. (**A**) Pie chart displaying the coordination geometries of zinc sites; (**B**) histogram reporting the occurrence of residues in the first coordination sphere of zinc ions; (**C**) distances between zinc ions and different donor atoms.

- *Per Geometry*, which provides statistics per each coordination geometry defined in FindGeo (15). By clicking on the geometry of interest, the user enters a page describing which metals were assigned that geometry in MetalPDB and how many different metal-binding patterns adopted that geometry for each metal;
- *Metal domains*, which provides an overview of the fractional occurrence of metal-binding domains in domain databases, in total and on a per-metal basis. For the SCOP and CATH databases, the per-metal statistics is further subdivided by domain class;
- *Per metal*, which enables two different kinds of analyses: coordination geometries or metal ligand distributions. In this page, the users selects one specific metal ion for which s/he wants to obtain statistics; then the desired analysis is selected by pressing a button at the bottom of the page. In the Geometries section, MetalPDB reports the occurrence of all regular coordination geometries (Figure 3A), the distribution of aminoacidic ligands for each geometry, and the number of different metal-binding patterns observed for the selected metal as a function of the coordination geometry. In the Ligands section, MetalPDB reports the statistics on the presence of aminoacidic or nucleic ligands in the coordination sphere of the selected metal (Figure 3B), the distribution of metal to donor atom distances (Figure 3C), and data on non-bonded interactions between aminoacidic ligands and other aminoacids of the protein (so-called second-sphere interactions);
- *Metals in enzymes*, which reports on the presence of metal sites in enzymes as well as on the occurrence of the different metal ions among the six EC classes and on the distribution of the six EC classes among metalloproteins on a per-metal basis (note that we include both catalytic and non-catalytic MFSs for any protein that has a EC number associated);
- *Metal substitutions in sites*, which reports on the distribution of the different metal ions replacing any given metal in all the sites (for example showing that the most common replacement for Ni is Zn, whereas for Mg it is Ca);

this statistics is derived from the comparison of equistructural groups.

All these pages are updated every time the database content is updated to the newest PDB release. Several of the statistics listed above, in particular those involving ligands and metal-binding patterns, address only metalloproteins.

CONCLUSIONS AND PERSPECTIVES

The number of structurally characterized metal-binding sites in biological macromolecules is still experiencing a significant growth. We have coped with this growth (64% in 6 years) by reviewing and improving the protocols for the construction of MetalPDB contents. In parallel, we expanded the options available to users for interacting with MetalPDB as well as the amount and complexity of precomputed structural and functional information displayed in the pages of MetalPDB. In the next releases of MetalPDB, we will continue to improve the functional information, also by enabling queries and statistics that target functional aspects directly. An important advancement is the functional annotation of individual MFSs, which is only partial at present. In the future development of MetalPDB we will work on increasing the coverage of annotated MFSs.

ACKNOWLEDGEMENTS

We gratefully acknowledge the technical help of Enrico Morelli. Maria Laura Vieri is thanked for the functional annotation of some copper-binding sites. We thank Dr Serena Lorenzini for an initial version of some of the statistics pages.

FUNDING

European Commission through the H2020 projects West-Life [675858]; EGI-ENGAGE [654142]. Funding for open access charge: European Commission [H2020].
Conflict of interest statement. None declared.

REFERENCES

1. Andreini,C., Bertini,I. and Rosato,A. (2009) Metalloproteomes: a bioinformatic approach. *Acc. Chem. Res.*, **42**, 1471–1479.
2. Andreini,C., Bertini,I., Cavallaro,G., Holliday,G.L. and Thornton,J.M. (2008) Metal ions in biological catalysis: from enzyme databases to general principles. *J. Biol. Inorg. Chem.*, **13**, 1205–1218.
3. Pechlaner,M. and Sigel,R.K.O. (2012) Characterization of metal ion-nucleic acid interactions in solution. *Met. Ions. Life Sci.*, **10**, 1–42.
4. Andreini,C., Cavallaro,G., Lorenzini,S. and Rosato,A. (2013) MetalPDB: a database of metal sites in biological macromolecular structures. *Nucleic Acids Res.*, **41**, D312–D319.
5. Rose,P.W., Prlic,A., Altunkaya,A., Bi,C., Bradley,A.R., Christie,C.H., Costanzo,L.D., Duarte,J.M., Dutta,S., Feng,Z. *et al.* (2017) The RCSB protein data bank: integrative view of protein, gene and 3D structural information. *Nucleic Acids Res.*, **45**, D271–D281.
6. Andreini,C., Bertini,I. and Cavallaro,G. (2011) Minimal functional sites allow a classification of zinc sites in proteins. *Plos ONE*, **10**, e26325.
7. Rosato,A., Valasatava,Y. and Andreini,C. (2016) Minimal functional sites in metalloproteins and their usage in strucutral bioinformatics. *Int. J. Mol. Sci.*, **17**, 671.
8. Valasatava,Y., Andreini,C. and Rosato,A. (2015) Hidden relationship between metalloproteins unveiled by structural comparison of their metal sites. *Sci. Rep.*, **5**, 9486.
9. Valasatava,Y., Rosato,A., Cavallaro,G. and Andreini,C. (2014) MetalS^3, a database-mining tool for the identification of structurally similar metal sites. *J. Biol. Inorg. Chem.*, **19**, 937–945.
10. Hutchinson,E.G. and Thornton,J.M. (1996) PROMOTIF–a program to identify and analyze structural motifs in proteins. *Protein Sci.*, **5**, 212–220.
11. Finn,R.D., Coggill,P., Eberhardt,R.Y., Eddy,S.R., Mistry,J., Mitchell,A.L., Potter,S.C., Punta,M., Qureshi,M., Sangrador-Vegas,A. *et al.* (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.*, **44**, D279–D285.
12. Maret,W. (2010) Metalloproteomics, metalloproteomes, and the annotation of metalloproteins. *Metallomics.*, **2**, 117–125.
13. Dawson,N.L., Lewis,T.E., Das,S., Lees,J.G., Lee,D., Ashford,P., Orengo,C.A. and Sillitoe,I. (2017) CATH: an expanded resource to predict protein function through structure and sequence. *Nucleic Acids Res.*, **45**, D289–D295.
14. Chandonia,J.M., Fox,N.K. and Brenner,S.E. (2017) SCOPe: manual curation and artifact removal in the structural classification of proteins - extended database. *J. Mol. Biol.*, **429**, 348–355.
15. Andreini,C., Cavallaro,G. and Lorenzini,S. (2012) FindGeo: a tool for determining metal coordination geometry. *Bioinformatics*, **28**, 1658–1660.
16. Nair,S.K., Calderone,T.L., Christianson,D.W. and Fierke,C.A. (1991) Altering the mouth of a hydrophobic pocket. Structure and kinetics of human carbonic anhydrase II mutants at residue Val-121. *J. Biol. Chem*, **266**, 17320–17325.
17. Zhu,D.W., Dahms,T., Willis,K., Szabo,A.G. and Lee,X. (1994) Crystallization and preliminary crystallographic studies of the crystals of the azurin Pseudomonas fluorescens. *Arch. Biochem. Biophys.*, **308**, 469–470.
18. Frausto da Silva,J.J.R. and Williams,R.J.P. (2001) *The Biological Chemistry of the Elements: the Inorganic Chemistry of Life*. Oxford University Press, NY.

5.2 MetalPredator version 2.0

5.2.1 Rationale

Minimal Functional Sites (MFSs), the core objects of MetalPDB [31], describe the local environment around the metal ion, independently of the larger context of the protein fold in which it is embedded. For zinc, almost 80% of the metal-binding sites found in different protein superfamilies have structures that can be classified into only 10 MFSs folds [87] (Figure 14). It is thus likely that MFS folds of metal-binding sites are less than those of entire proteins. Using the most recent CATH classification [88] the latest protein structure with a novel fold deposited in PDB dates back to 2009, and that belonging to a novel protein superfamily dates back to 2010. We thereby expect that the majority of the folds of metal-binding sites is already represented in MetalPDB. We also showed that the sequence of MFSs is generally more conserved than that of entire proteins (unpublished data). We thus expect that MFS sequence profiles are able to identify metal-binding protein sequences with a higher sensitivity than sequence profiles of entire proteins or of protein domains. Furthermore, the analysis of zinc-binding sites shows that they can be seen as composed of recurrent structural modules which combine each other in different ways to generate different sites. Figure 15 shows some examples of the occurrence of *β-hairpin* in zinc-binding sites. The order of these structural modules in the protein sequence can also vary, as shown in Figure 16. Sequence profiles generated from the alignment of such modules (metal-binding motifs, hereafter) can be used to identify novel combinations of modules, i.e., novel types of metal-binding site in protein sequences.

In this work, we developed the second version of *MetalPredator*, to predict iron- (heme and ion), zinc- and copper- binding sites in protein sequence(s) at the whole proteome scale. The tool integrates an existing domain-based approach [66] with metal-binding motifs derived from MFSs in proteins structures. MetalPredator uniquely combines global and local searches to define whether a protein is a potential metal-binding protein.

Figure 14. *An MFS fold found in 61 distinct superfamilies (only nine shown)*

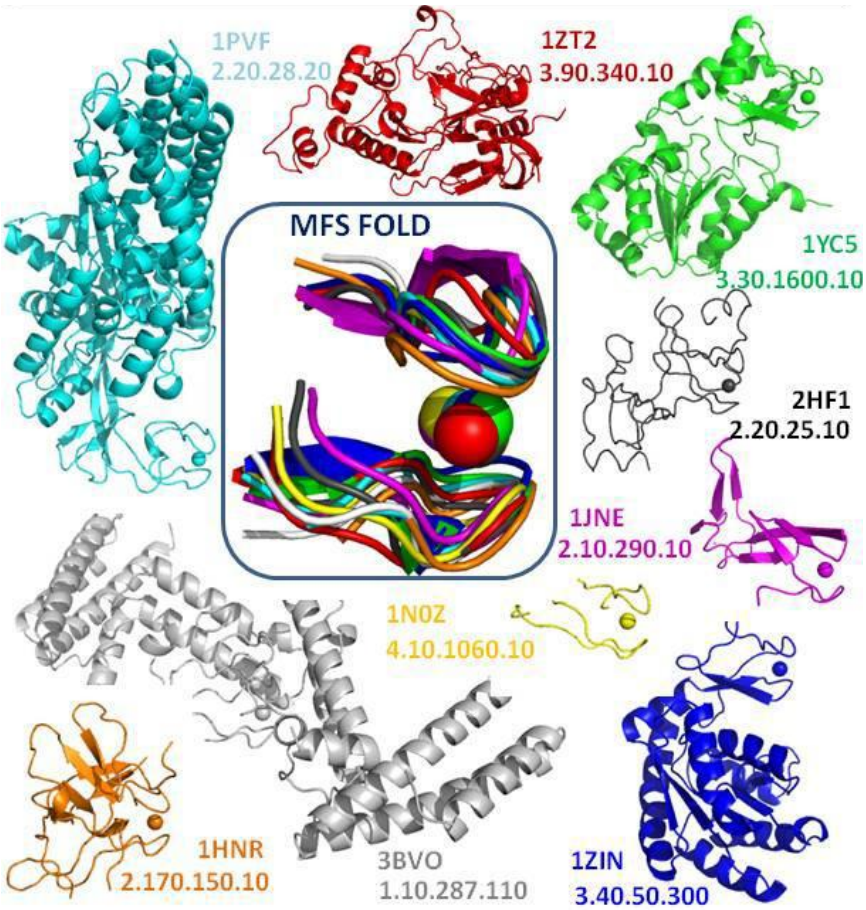


Figure 15. *A recurring β-hairpin module (red) in zinc MFSs*

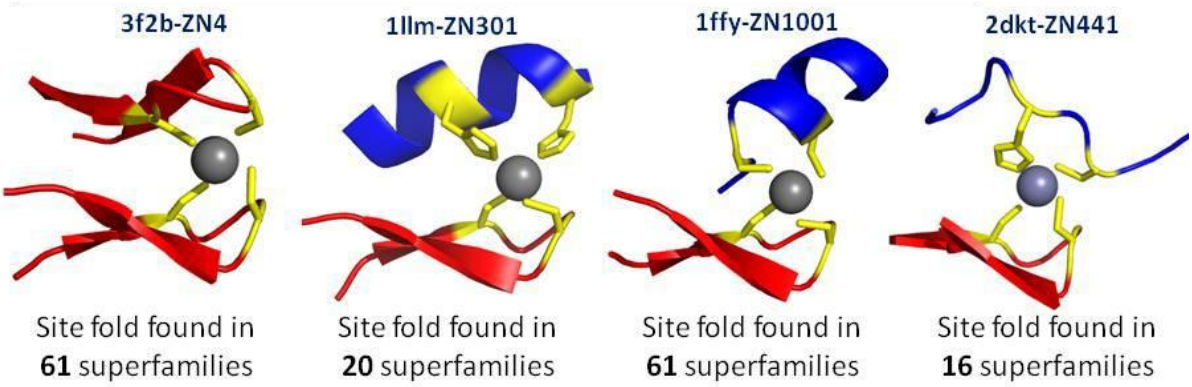
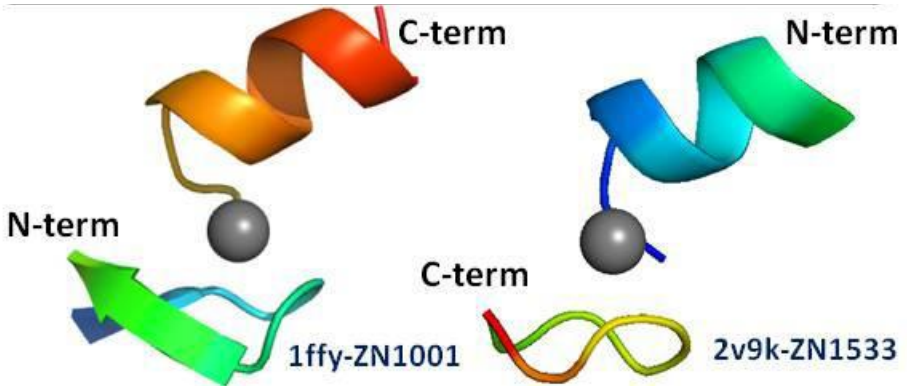


Figure 16. *Two swapped modules within the same MFS fold (ranbow colored from N- to terminal)*



5.2.2 MetalPredator overview

MetalPredator uses two libraries of Hidden Markov Model profiles to identify metal-binding sites in protein sequences, i.e. (1) Pfam[47] domains and (2) metal-binding motifs. Library (1) was built as described in [66]. It contains profiles of zinc- copper and iron-binding Pfam domains. Metal-binding motifs are defined by splitting the Minimal Functional Sites (MFSs) stored in MetalPDB[31] into fragments. Each fragment is a continuous stretch of protein sequence containing at least one metal ligand. To build the library (2) each zinc- copper- and iron-binding MFS in MetalPDB was searched into UniRef50 [68] using PSI-Blast [67]. All the hits with conserved ligands were used to build a sequence profile. From this profile we extracted the profiles of the distinct fragments corresponding to the MFS(s) in the initial input sequence.

MetalPredator uses the hmmscan tool [70] to match every input sequence to the profiles contained in the libraries. An input sequence is identified as a potential zinc-, copper- or iron-binding protein if at least one of these conditions applies:

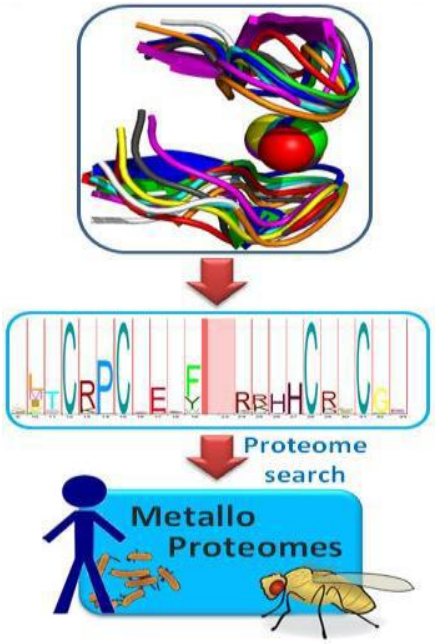
(A) The profile of a Pfam domain with associated ligands (library 1) matches the sequence with an e-value lower than 10^{-5} and ligands are conserved in the sequence.

(B) The profile of a domain with no information on ligands available (library 1) matches the sequence with an e-value lower than 10^{-7}.

(C) All fragment profiles of a given MFS (library 2) match the sequence with an e-value lower than 10^{-3} and the corresponding ligands are conserved in the sequence (Figure 17).

(D) At least one fragment profile of a given MFS (library 2) matches the sequence with an e-value lower than 10^{-3} and the corresponding fragment ligands are conserved in the sequence (Figure 17).

Figure 17. *Metalloproteome search based on MFS profiles*



5.2.3 Performances of MetalPredator

MetalPredator was designed to predict zinc-, copper- and iron-binding (iron ions, heme, and iron-sulfur clusters) sites in protein sequences. For the calibration of parameters, we used, for each metal two datasets of sequences (positives and negatives), taken from a subset of the Protein Data Bank [28] filtered at a sequence identity level of 25% (PDB25) (Table 6).

We assessed performances of the method for each metal-cofactor (Table 7). To avoid overfitting, we assessed MetalPredator by using a leave-one-out cross-validation (LOOCV) approach on the entire PDB25. In LOOCV each training set is created by taking all the samples except one, and the test set is the sample left out. The procedure is repeated by creating as many training and test sets as are the samples available. For each cofactor, we further test MetalPredator against datasets of the other cofactors in order to establish how many profiles are not specific to a given metal ion. These results are reported in Table 8.

Table 6. *Dimension of Positive and Negative datasets used to calibrate the method*

Metal cofactor	**Negative dataset**	**Positive dataset**
Iron ion	9835	298
Iron (heme)	9860	273
Iron-sulfur cluster	2707	163
Zinc ion	7140	1822
Copper ion	9368	124

Table 7. *Prediction performances of MetalPredator*

Metal cofactor	**Sensitivity (%)**	**Specificity (%)**	**Accurancy (%) ***
Iron ion	72,8	92,8	82
Iron (heme)	94,1	97,5	96
Iron-sulfur cluster	86,0	82,5	-
Zinc ion	74,6	86,0	81
Copper ion	74,1	95,6	84

* The dataset of negatives is more big than the positive, so the accuracy was calculated using the formule:
VP + [VN/(tot neg/tot pos)] / VP + FN + [VN/(tot neg/tot pos)] + [FP/(tot neg/tot pos)]

Table 8. *Test of MetalPredator on negative datasets of different metal ions*

	Fe ion	**Fe heme**	**Zn ion**	**Cu ion**	**Total**
Zn ion	155	45	-	82	1.822
Mg ion	67	27	251	86	2.168
Ca ion	62	47	177	81	1795
Na ion	65	38	183	64	1539
Mn ion	98	14	144	16	563
Ni ion	62	8	95	24	399
Co ion	59	13	83	14	246
Fe ion	-	15	32	10	298
Fe heme	12	-	10	8	273
Fe-S	42	3	17	7	163
Cu ion	7	11	25	-	124
Mo ion	0	3	3	5	20

5.2.4 The human iron-proteome

We used MetalPredator 2.0 to carry out a systematic prediction of iron-binding proteins encoded in the human genome. In total, we identified 398 human genes whose protein products interact with iron, which correspond to about 2% of the all human genes. Of these, 139 genes express proteins binding individual iron ions, 192 express proteins binding heme and 70 express proteins binding iron-sulfur clusters. Among the identified iron-binding proteins only for 105 proteins is available a 3D structure in the iron-bound form, while for 76 proteins is available a structure of a close homolog (sequence identity at last 50%) of the human protein in the iron-bound form.

Metallomics

PAPER

View Article Online
View Journal | View Issue

Cite this: *Metallomics*, 2018, **10**, 1223

The human iron-proteome†

Claudia Andreini,[ab] Valeria Putignano,[a] Antonio Rosato[ab] and Lucia Banci*[ab]

Organisms from all kingdoms of life use iron-proteins in a multitude of functional processes. We applied a bioinformatics approach to investigate the human portfolio of iron-proteins. We separated iron-proteins based on the chemical nature of their metal-containing cofactors: individual iron ions, heme cofactors and iron–sulfur clusters. We found that about 2% of human genes encode an iron-protein. Of these, 35% are proteins binding individual iron ions, 48% are heme-binding proteins and 17% are iron–sulfur proteins. More than half of the human iron-proteins have a catalytic function. Indeed, we predict that 6.5% of all human enzymes are iron-dependent. This percentage is quite different for the various enzyme classes. Human oxidoreductases feature the largest fraction of iron-dependent family members (about 37%). The distribution of iron proteins in the various cellular compartments is uneven. In particular, the mitochondrion and the endoplasmic reticulum are enriched in iron-proteins with respect to the average content of the cell. Finally, we observed that genes encoding iron-proteins are more frequently associated to pathologies than the all other human genes on average. The present research provides an extensive overview of iron usage by the human proteome, and highlights several specific features of the physiological role of iron ions in human cells.

Received 20th June 2018,
Accepted 3rd July 2018

DOI: 10.1039/c8mt00146d

rsc.li/metallomics

Significance to metallomics

Iron is one of the most ancient and abundant metal ions in living organisms: it participates in fundamental biological processes, such as photosynthesis, and respiration. It is an essential metal ion for humans. Here, we applied a bioinformatics approach to predict the entire set of human proteins that use iron as cofactor. We found that about 2% of human genes encode an iron-protein. In particular, 35% are proteins binding individual iron ions, 48% are heme-binding proteins and 17% are iron–sulfur proteins. Most of these proteins are enzymes: 37% of the human oxidoreductases need an iron ion to perform their catalytic mechanisms. The analysis of the subcellular location highlighted that some organelles are enriched in iron-proteins, in particular about 7% of the proteins localized in the endoplasmic reticulum and in the mitochondrion bind iron. Finally, our data show that mutations in genes encoding iron-binding proteins are more likely to be associated with pathology than all human genes on average.

Introduction

During evolution, organisms have selected some of the available elements from the environment to catalyze physiological reactions. Consequently, some metal ions became essential to life. Iron is one of the most ancient and abundant transition metal ions in living organisms,[1,2] as it was highly available as ferrous ion in the early days of terrestrial life.[3] Iron is essential to all forms of life and participates in fundamental biological processes, such as photosynthesis, respiration and nitrogen fixation.[4,5] In cells, it is normally found in the +2 (ferrous) and/or +3 (ferric) oxidation states. Higher oxidation states may be generated transiently in the course of the catalytic cycle of enzymatic reactions. Besides individual iron ions, proteins can bind also iron-containing cofactors, such as heme or iron–sulfur clusters.[6–8] Heme is one of the most versatile prosthetic groups in metalloproteins. The porphyrin constituting the heme group can be of several types, including *e.g.* heme a, heme b, and heme c. The heme proteins that transfer electrons mainly belong to the cytochromes class, and may contain one or several heme groups; globins are heme-containing proteins involved in dioxygen binding and/or transport; other heme proteins serve as biological sensors for oxidative stress. The broad range of possible reactions occurring at the heme center is mainly based on the ability of the heme iron to coordinate small molecules like CO, NO, and O_2. The protein matrix can modulate the affinity towards the different exogenous ligands. Iron–sulfur clusters contain two or more iron ions bridged by sulfide ions. Each iron ion is tetracoordinated, with its coordination sphere typically completed by the sulfur

[a] *Magnetic Resonance Center, University of Florence, Via Luigi Sacconi 6, Sesto Fiorentino 50019, Italy. E-mail: banci@cerm.unifi.it Fax: +39 055 4574253; Tel: +39 055 4574273*
[b] *Department of Chemistry, University of Florence, Sesto Fiorentino 50019, Italy*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c8mt00146d

or nitrogen atoms of cysteine and histidine side chains, respectively.[9] The metal site of rubredoxin, which contains a single iron ion coordinated by four cysteines, is generally classified as the simplest unit of iron–sulfur clusters. Iron–sulfur clusters are among the most versatile inorganic cofactors.[5] They are involved in a plethora of functional processes, including aerobic as well as anaerobic respiration, regulation of gene expression, amino acid and nucleotide metabolism, DNA modification and repair and tRNA modification.

Heme and iron–sulfur clusters are cofactors featuring a high chemical complexity. Therefore, their biosynthesis as well as the biosynthesis of the final holo-proteins containing these cofactors involve a significant number of different protein components, some of which are iron-binding proteins. In the human cell, these biosynthetic processes have multiple pathways, related also to cellular compartmentalization. Nevertheless, some components may move across different compartments; furthermore, the various pathways can communicate with one another *via* the exchange of biosynthetic intermediates.

While iron is essential for life, it can catalyze the formation of potentially toxic reactive oxygen species (ROS). This process is unavoidable in the present oxygen-rich environment, and iron and ROS are increasingly recognized as important initiators and mediators of cell death in various organisms as well as in pathological conditions in humans.[10] Therefore, biological systems must control iron metabolism by providing the adequate amount of iron for proper cellular function while limiting iron toxicity.[11,12] Iron has also a role in pathogen virulence. The growth of microbial pathogens within the host usually requires iron as an essential nutrient.[13,14] Heme-containing proteins, such as hemoglobin, and transferrin are the preferential iron sources for human pathogens.[15,16] Therefore, another crucial reason for the cell to maintain a strict control on iron homeostasis is to restrict its access by pathogens.

In this paper, we carried out a systematic prediction of iron-binding proteins encoded in the human genome, extending our previous analysis on iron–sulfur proteins.[17] By integrating this prediction with information on heme and individual iron ions, we achieved a complete landscape of the iron handling by proteins in human, thus providing a framework for the understanding of physiological iron metabolism and of its dysfunction in diseases.

Results

Iron binding by human proteins and their coordination spheres

We analysed iron usage by human proteome *via* three different possible modes of binding: as individual iron ions, as iron-containing heme cofactors and as iron–sulfur clusters. In total, we identified 398 human genes whose protein products interact with iron (iron-proteins hereafter), *i.e.* about 2% of the human genes. Of these, 139 genes express proteins binding individual iron ions (Table S1, ESI†), 192 express proteins binding heme (Table S2, ESI†) and 70[17] express proteins binding iron–sulfur clusters (Table S3, ESI†).

The coordination spheres of the three different iron-containing cofactors are quite diverse; we refer to the pattern of the protein residues coordinating the iron ion(s) of the cofactor as the iron-binding pattern (IBP). The IBP is a regular expression defined by the identity of the amino acids coordinating the metal and by their spacing along the protein sequence (*e.g.* $CX_4CX_{25}C$). Thus, the coordination sphere of each iron ion corresponds to a single IBP.

In IBPs of human iron-proteins binding individual iron ions, histidine is by far the most common residue. His is present in 94% of these IBPs, each of which contains on average two His (Fig. 1). Aspartate, glutamate and tyrosine are found in 53%, 30% and 10% of the identified patterns, respectively. On average, only one Asp and one Tyr are found in each IBP, whereas there can be one (such as in most iron-dependent enzymes) or two (such as in ferritins) Glu residues. All iron–sulfur binding proteins use on average three-four cysteines to coordinate the cluster. Cys is absolutely required in the IBPs of these proteins. In particular, in human iron–sulfur proteins the coordination sphere of the Fe_4S_4 clusters is always and only composed by cysteines whereas the IBPs of Fe_2S_2 clusters sometimes (37% of Fe_2S_2 IBPs) include one or two His residues. In human heme-binding proteins, IBPs commonly contain one or two His with the exception of catalytic heme sites (such as in cytochrome P450) where Cys is more common (83% of IBPs).

The function of the metal cofactor within the protein is correlated also to the number of coordinating residues provided by the protein (*i.e.* the number of residues in the IBP).

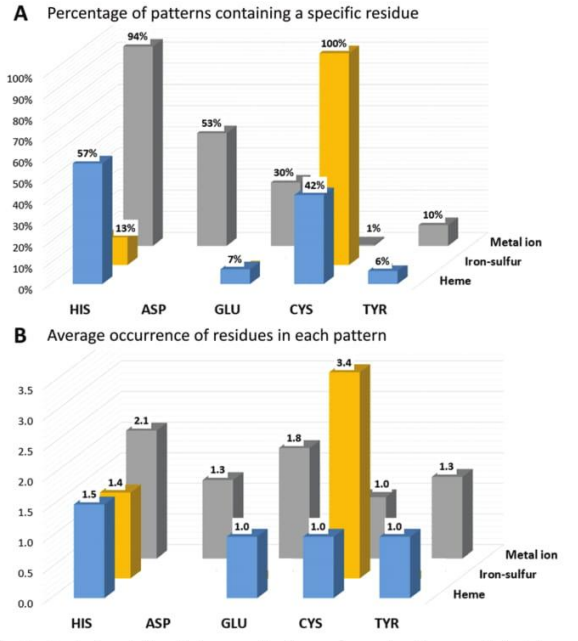


Fig. 1 Analysis of the first coordination sphere for the predicted iron-proteins; (A) percentage of patterns containing a specific residue for different iron cofactor types. (B) Average occurrence of a specific residue within patterns, for each iron cofactor.

Indeed, the coordination sphere of the metal ion is not always completed by atoms of the protein. 64% of the sites that bind individual iron ions contain three protein residues in the IBP, whereas the others contain four protein residues. Similarly, most of the iron ions in heme cofactors have only one ligand provided by the protein (about 58%), which allows the substrate to occupy the second heme axial position. The remaining 42% heme sites have two coordinating residues provided by the protein. In iron–sulfur proteins, the most common number of protein ligands is 4; however, all the iron–sulfur clusters that perform a catalytic function have only three Cys ligands in the IBP. It is thus evident that there is a trend for human iron-proteins to have a lower number of residues in their IBPs when the metal-binding site performs a catalytic function, in order to allow the iron ion to coordinate directly to the substrate as already observed for other metal containing proteins.[18]

Subcellular localization of human iron-proteins

We then analysed the subcellular localization of the human iron-proteins identified through our search (Tables S4–S6, ESI†). This information is not available for 94 proteins (37 binding individual iron ions, 10 binding iron–sulfur clusters, and 47 binding hemes), which were thus ignored for this analysis. Various proteins are present in more than one compartment, and thus were included in the statistics of each relevant organelle. Fig. 2 summarizes the distribution of the different types of iron-proteins within each cellular compartment and reports the fraction of iron-proteins with respect to the total number of proteins localized in each compartment (percentages within parenthesis). It appears that two subcellular locations stand out for their enrichment in iron-proteins: the mitochondrion and the endoplasmic reticulum.

Our dataset (iron-proteins for which cellular localization is known) is composed by 45% heme-binding proteins, 34% proteins binding individual iron ions, and 21% proteins binding iron–sulfur clusters. From Fig. 2, we can readily identify compartments that differ appreciably in the distribution of the types of iron-proteins. The nucleus is highly depleted of heme-binding proteins, whereas it features a relatively high number of proteins binding individual iron ions. On the other hand, the mitochondrion is the compartment most enriched in iron–sulfur proteins, with respect to both the two other types, whereas the endosome is mostly enriched in heme-binding proteins and does not contain any iron–sulfur protein. In addition, the endoplasmic reticulum is enriched in heme-binding proteins and depleted in iron–sulfur proteins. The distribution of the three types of iron-proteins in the cytoplasm closely resembles that of the overall dataset. It should be noted that in this respect, we are referring to the number of proteins and not to their relative quantity, which depends on their expression levels. We did not analyze such levels in this work.

The mitochondrion and the endoplasmic reticulum are the compartments with the largest percentage of iron-proteins. As mentioned, the mitochondrion is significantly enriched in iron–sulfur proteins (about 2.5 times the average fraction for the whole cell), whereas the endoplasmic reticulum is enriched in heme-binding proteins (1.6 times the cell average). The nucleus is the only compartment where proteins binding individual iron ions are the majority of iron-proteins (1.7 times the cell average).

Functional roles

Fig. 3 shows the functional roles of sites binding iron and iron-containing cofactors in human proteins (Tables S4–S6, ESI†). This information is not available for 24 proteins (14 binding iron–sulfur clusters, and 10 binding heme), which were thus ignored for this analysis. It appears that sites binding heme or individual iron ions most commonly have a catalytic role, *i.e.* are directly involved in enzymatic mechanisms. This is also the most common role for the entire set of iron-proteins, partly due to the low number of iron–sulfur proteins. For sites binding individual iron ions the only other relevant function is its use as a substrate, *i.e.* in storage and transport processes

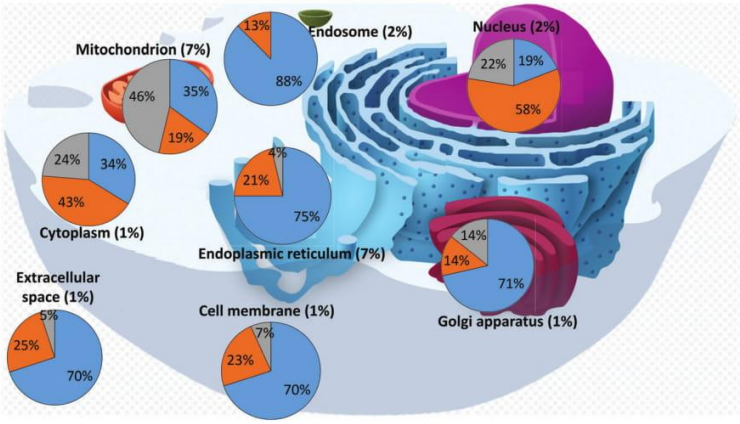


Fig. 2 Distribution of iron-proteins in different cellular organelles of the human cell (heme-proteins: blue; iron–sulfur proteins: grey; individual iron ions: orange).

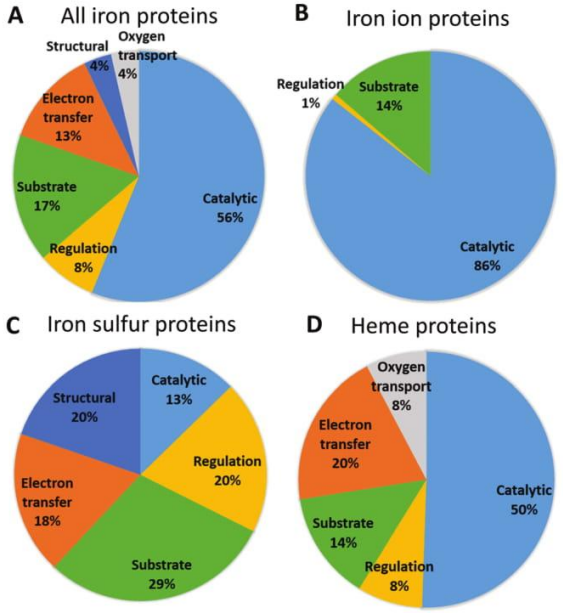


Fig. 3 Distribution of the functions of the iron centers for different iron cofactor types.

Table 1 Number of genes coding for iron-proteins in the endoplasmic reticulum, nucleus and mitochondrion. Note that the same gene can contribute to more than one process in each compartment. Processes are taken from the GO annotations of all iron-protein genes

	All	iron_ion	iron_heme	iron_sulfur
Endoplasmic reticulum				
Drug metabolism	14	0	14	0
Peptidyl amino acid hydroxylation	6	6	0	0
Lipid metabolic process	43	5	38	0
Cell proliferation	12	4	8	0
Response_to_stress	9	0	9	0
Vitamin metabolism	8	0	8	0
Xenobiotic metabolic process	20	0	20	0
Nucleus				
Cell death/apoptotic process	20	10	5	5
Gene expression	46	33	9	4
Cell proliferation	20	11	5	4
Peptidyl amino acid hydroxylation	8	8	0	0
Response to stress	25	9	6	10
Mitochondrion				
Cell death/apoptotic process	13	4	5	4
Iron ion homeostasis	11	4	4	3
Iron sulfur cluster biosynthesis	6	0	0	6
Cellular respiration	18	1	7	10
Response to drug	9	1	5	3
Response to stress	16	3	5	8

(this classification of sites is taken from the MetalPDB database[9]). Heme-binding sites have the largest variety of functional roles, among which electron transfer is the second most common. As it is well known, human heme-binding proteins also play a crucial role in the transport of molecular dioxygen and in sensing, particularly of small gaseous molecules such as NO, leading to a regulatory function. Heme-binding proteins associated with a substrate function (*i.e.* when the heme cofactor is the target/substrate of the protein) are involved in the biosynthesis, transport and degradation of the heme cofactor. This may be linked also to the fact that there are as many as seven different types of heme cofactors in human heme-binding proteins (heme a, b, c, d, i, o, m). While the most common type is heme b, occurring in 90% of the heme-proteins, the synthesis of all the other heme types requires the action of specific enzymes that modify the cofactor and/or the protein binding it (*e.g.* cytochrome *c*[19]).[20,21] The most common role for iron–sulfur proteins is transport, biosynthesis and insertion into the final target proteins of the clusters themselves (tagged as substrate).[22–26] This is the result of both the chemical complexity of the iron-containing clusters, thus requiring elaborate biosynthetic and degradation pathways, and the potential toxicity of free iron ions. The second most common roles for iron–sulfur proteins are structural and regulatory. The role of iron–sulfur clusters in several DNA- and RNA-binding proteins is not completely understood, in particular for the many systems involved in DNA repair, where the presence of the cluster could be instrumental to detect lesions. Curiously, sites performing electron transfer are less common.

We then checked whether there is a relationship between cellular localization and protein function in order to rationalize the patterns reported in Fig. 2. To do this we examined the lists of the iron-proteins localized to the various compartments and identified all the processes, as defined by the Gene Ontology (GO[27,28]), associated with the corresponding genes. Seven processes involve 81% of the genes coding for iron-proteins localized to the endoplasmic reticulum (Table 1). The process involving more iron-proteins is lipid metabolism, which is a key cellular role played by cytochromes P450; only one tenth of the genes involved in lipid metabolism codes for proteins binding individual iron ions. Xenobiotic metabolic process and drug metabolism are common processes which involve exclusively heme-binding proteins and are essentially associated to cytochromes P450, which are involved in the modification of exogenous molecules, from drugs to pollutants. Proteins binding individual iron ions are involved in different pathways, such as peptidyl amino acid hydroxylation. These pathways do not involve any heme-binding protein. Overall, 92% of the iron-proteins localized to the endoplasmic reticulum are oxidoreductases, as directly observed from their Enzyme Commission (EC) numbers, and these are either members of the cytochrome P450 family (heme-containing enzymes) or iron-dependent hydroxylases (typically harboring two iron ions in their active site). The functional role of the iron-proteins in the endoplasmic reticulum is thus tightly linked to their catalytic activity, most commonly in biosynthetic or metabolic processes.

In the nucleus, 5 processes involve about 89% of the iron-proteins present in this cell compartment. Gene expression is the process associated to most of these proteins, because several genes encode iron-proteins involved in the regulation of transcription *e.g.* through DNA binding or histone modification. Many iron-proteins in the nucleus are also involved in

response to stress, for instance by repairing damaged DNA, in apoptosis[17] and in cell proliferation. About half of the nuclear iron-enzymes are oxydoreductases; transferases and hydrolases are relatively common.

In the mitochondrion, 6 processes involve about 63% of all iron-proteins within this cellular compartment. The process involving the largest number of iron-proteins is cellular respiration, which leverages both heme-binding and iron–sulfur proteins (6 *vs.* 10 genes, respectively). Other processes involving more than 10 genes are cell death, iron ion homeostasis and response to stress (which is mainly response to oxidative stress), half of which are iron–sulfur proteins. The biosynthesis of iron–sulfur clusters comprises genes encoding require iron–sulfur proteins. At the functional level, the observed enrichment of the mitochondrion in iron–sulfur proteins (Fig. 2) is largely accounted for by the involvement of these proteins in the respiratory chain, in stress response and in the assembly of iron–sulfur clusters themselves. For the latter, the clusters are transiently bound by various proteins along the biosynthetic pathway, also depending upon the final target for cluster insertion.[25,26,29] The electron transfer capabilities of iron–sulfur proteins are important but not the only determinant of the higher abundance in the mitochondrion of iron–sulfur proteins with respect to all iron-proteins.

Uncharacterized putative human iron-proteins

Our analysis identified several proteins that had not been described in the literature as binding iron or iron-containing cofactors. In particular, Retinoid-related Orphan Receptors-alpha, beta and gamma (RORα, RORβ, and RORγ, hereafter) were predicted to have a heme-binding site similar to that found in REV-ERBα and REV-ERBβ. The REV-ERB family binds heme with two axial ligands: one His and one Cys.[30] The sequence alignment of these two families (Fig. S1, ESI†) clearly shows that the His ligand is strictly conserved also in the ROR family whereas the Cys ligand is not. However, the superimposition of the heme-containing 3D structure of REV-ERBβ (PDB code 3CQV[30]) with the experimental structures of RORα, RORβ and RORγ (PDB codes 1N83,[31] 1NQ7,[32] 4WLB,[33] respectively) shows that the latter contain a Cys (Cys323, Cys262 and Cys320, respectively) that is essentially in the same position as the heme-binding Cys384 of REV-ERBβ (Fig. 4A). A small rearrangement of the side chains of the Cys residues would bring their Sγ atoms at a distance from the iron ion compatible with the formation of a coordination bond. This Cys corresponds to a strictly conserved position in the multiple sequence alignment of the ROR family (Fig. S1, ESI†). Furthermore, the cavities of the 3D structures of ROR are sterically compatible with the binding of a heme molecule and the regions in contact with the cofactor have a high sequence similarity with the REV-ERB family. Another new putative heme-binding protein is the extracellular matrix protein FRAS1. This protein is in the plasma membrane: it has a very long region exposed in the extracellular matrix and a short cytoplasmatic tail. We identified three putative heme-binding sites in the extracellular part. We predicted the occurrence of a site with two potential axial

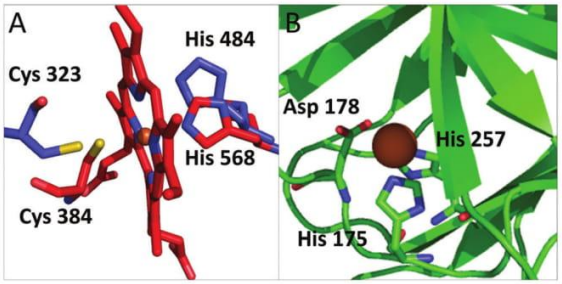


Fig. 4 (A) Superposition of RORα (pdb code: 1n83, in blue) and REV-ERB (pdb code: 3cqv, in red). Only the relative positions of the putative ligands of RORα and the iron ligands of REV-ERB are reported. The side chain of Cys 323 is rotated to bring it closer to the heme iron. In this configuration the distance between the potential sulfur donor and the iron ion is 3.4 Å. (B) Putative iron-binding site in the structural model of HSPB1-associated protein 1.

ligands (His2080 and His3301) whereas for the other two sites, we predicted only one ligand, *i.e.* His1799 and His1945, respectively. The structure of this protein is not available and we were not able to build a 3D structural model, which would have allowed us to evaluate the possible geometrical features of the three predicted sites. The HSPB1-associated protein 1 is another potential iron-binding protein which could bind a single iron ion *via* its residues His175, Asp177 and His257; all these three residues are highly conserved in the protein family. For this protein we could identify a suitable template in the PDB for 3D structural prediction by homology modeling: the Hypoxia-inducible factor 1-alpha inhibitor which has a sequence identity to human HSPB1-associated protein 1 as high as 26%, and contains a site binding a single iron ion. The structural model in Fig. 4B, shows that the predicted ligands of HSPB1-associated protein 1 have the proper spatial configuration to bind an iron ion. Finally, we predicted as putative heme-binding protein the phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase 2. A structure as well as a suitable 3D template for the putative heme-binding region of this protein are not available. This prediction, however, appears less reliable than the previous ones.

Pathogenic alterations associated to human iron-proteins

To assess the impact of the iron-proteome on the human health, we investigated how often defects or mutations affecting genes encoding iron-proteins are associated to pathologies (Tables S4–S6, ESI†). We analysed only proteins in the Swiss-Prot database (Reviewed proteins)[34] and excluded those from the trEMBL database, which are just predicted and do not have mutational studies associated. Thus, we took into account 385 proteins (137 binding individual iron ions, 178 binding heme, and 70 binding iron–sulfur clusters). Of these, 148 are related to one or more pathogenic mutations or alterations, corresponding to about 38% of the total. Interestingly, if we consider the different types of iron sites, we found that more than half of the identified iron–sulfur proteins are involved in pathologies (37/70 corresponding to 53%). For proteins binding individual iron ions or heme cofactors, the percentage of proteins

Table 2 Number of proteins associated to at least one pathology in UniProt and their ratio with respect to the total number of iron proteins in each cellular compartment, and compared with the data for all human proteins. The percentage of disease-related proteins is in parentheses

	Heme	Individual iron-ions	Iron–sulfur clusters	Total iron-proteins	All human proteins
Cytoplasm	13/27 (48%)	10/34 (29%)	8/19 (42%)	31/80 (39%)	1413/5569 (25%)
Endoplasmic reticulum	15/60 (25%)	9/17 (53%)	0/3 (0%)	24/80 (30%)	362/1163 (31%)
Mitochondrion	20/28 (72%)	5/15 (33%)	23/37 (62%)	48/80 (60%)	420/1174 (36%)
Nucleus	7/17 (41%)	10/52 (19%)	11/20 (55%)	28/89 (31%)	1180/5389 (22%)

associated to pathologies is 31% (*i.e.* 43/137) and 38% (*i.e.* 68/178), respectively. As of January 2018, the total number of human proteins in the Swiss-Prot database was 20259. Of these, 4014 are associated to pathogenic mutations, corresponding to about 20% of the dataset. It thus appears that on average defects or mutations affecting genes encoding iron-proteins are more commonly associated to pathologies than all the other genes.

In Table 2 we broke down the cumulative data reported in the previous paragraph for the whole human cell by looking at specific compartments. In particular, we took into consideration the compartments with the highest number of iron-proteins. In the mitochondrion, 36% of all proteins are associated to pathologies, whereas as many as 60% of mitochondrial iron-proteins are disease-related, with the main contribution of heme-proteins and iron–sulfur proteins. Similarly, in the cytoplasm and in the nucleus, heme-proteins and iron–sulfur proteins are more commonly associated to pathologies than all other human genes (Table 2).

Discussion

398 human genes encode iron-proteins, which correspond to about 2% of all human genes. This number should be regarded as a lower limit because within our approach to the identification of iron-proteins false positives (*i.e.* proteins that do not bind iron but are predicted to do so) are quite unlikely to occur. This is due to the fact that we rely significantly on the known 3D structures of iron-proteins, while in the absence of structural data we scan the literature for supporting evidence. On the other, it is possible that we did not detect completely uncharacterized iron-proteins, especially if they are membrane-associated. Therefore, this number (398) should be taken as a lower limit even if we foresee that the actual number should not be much different.

Of the 398 human iron-proteins, 48% are heme-binding proteins, 35% are proteins binding individual iron ions and 17% are iron–sulfur proteins. The intracellular distribution of these proteins is uneven, with some organelles containing a larger share of iron-proteins than others do. In particular, 7% of all the proteins localized in the endoplasmic reticulum and in the mitochondrion are iron-proteins. Thus these two organelles are significantly enriched (in comparative terms) in iron-proteins with respect to the average of the entire human cell (2%, as mentioned above). Within heme-binding proteins, 90% bind heme b and 61% are membrane-associated.

The three types of iron-proteins feature highly diverse preferences in the coordination sphere of the bound iron ions (*i.e.* IBPs). Cys is always present in the IBPs of iron–sulfur proteins, whereas it is practically absent from the coordination sphere of individual iron ions. Conversely, His, which is nearly always present in the IBPs of proteins binding individual iron ions, is observed rarely in the IBPs of iron–sulfur proteins. Asp is the second most common ligand in proteins binding individual iron ions. Heme-proteins have a similar preference for His and Cys in their IBPs. Cys is particularly common in the IBPs of heme-proteins that have catalytic function. This is presumably linked to the role of Cys in promoting the heterolytic breakage of the O–O bond of the iron-bound peroxide intermediate that forms along the catalytic cycle of cytochromes P450 or of nitric oxide synthase.[35–37] This feature is independent of the overall protein fold, and is defined by the coordination chemistry properties of the sites.

6.5% of the human enzymes are iron-proteins. Unsurprisingly, this percentage is not the same for all enzyme classes. In particular, 37% of human oxidoreductases use a catalytic iron ion. 56% of all human iron-proteins have a catalytic function (Fig. 3). Proteins that bind individual iron ions mainly represent them: 86% of these proteins (119 out of 139) are iron-dependent enzymes. The large majority of these enzymes are oxidoreductases, in particular dioxygenases, where the iron ion is directly involved in the transfer of electron from/to the substrate. Also, about half of the heme-sites in the human proteins have a catalytic function. These enzymes are primarily members of the human cytochrome P450 family, whose isoforms are significantly differentiated in terms of expression but have typically broad and overlapping substrate specificities.

Iron-binding enzymes are commonly located in the nucleus and cytoplasm, followed by the mitochondrion and endoplasmic reticulum. The latter features the highest number of heme-binding proteins as it is the most common localization for cytochromes P450. Consistently with this, we observed that processes such as drug metabolism, lipid metabolism or xenobiotic stimulus are the most common processes associated with iron-proteins localized to the endoplasmic reticulum (Table 1). In the mitochondrion, 63% of all iron-proteins are involved in only 6 processes; the process involving the largest number of iron-proteins is respiration, which leverages both heme-binding and iron–sulfur proteins. The mitochondrion is the most likely localization for iron–sulfur proteins (Fig. 2), whose primary processes within this compartment are, besides respiration, the biosynthesis of iron–sulfur clusters and the response to oxidative stress. The biosynthesis of iron–sulfur clusters is among the most common functional roles of iron–sulfur proteins at the level of the whole cell,[17,38] owing to the chemical

complexity of this group of cofactors. Within the nucleus, iron-proteins are largely involved in various aspects of the regulation of protein expression, such as histone modification. In addition, also DNA binding, DNA biosynthesis and DNA replication involve several iron-proteins, especially iron–sulfur proteins.

We identified three human members of the retinoid-related orphan receptor (ROR) family as potentially harbouring a heme-binding site similar to those observed in proteins of the REV-ERB family. In the absence of experimental evidence in the literature, our hypothesis is supported by the strict conservation of the two potential heme ligands. The experimental structures of RORα, RORβ, and RORγ, feature a His and a Cys residue in a spatial position corresponding to His and Cys ligands of iron in REV-ERBβ. Another putative human iron-binding protein is the HSPB1-associated protein 1. A structural model of this proteins shows that the reciprocal position in 3D space of the putative ligands is completely consistent with our prediction (Fig. 4).

As an important aspect of the present study, we analysed how many pathologies are associated to human genes encoding iron-proteins, based on the occurrence of disease-associated mutations reported in the Swiss-Prot database. The percentage of pathologies associated to genes encoding iron-proteins is almost 40%, which is higher than the percentage of pathologies associated to all human genes (about 20%). In practice, two genes out of 10 are associated with pathogenic mutations in the human genome, whereas this percentage is essentially doubled if we take into account specifically the genes encoding iron-proteins. Interestingly, this percentage peaks at 72% for all heme-binding proteins in the mitochondrion.

In summary, this work provided an extensive overview of iron usage by human proteins, spanning from iron coordination properties to biochemical/cellular function and compartmentalization, and addressing the interplay between these aspects. We observed that the distribution of the type of iron cofactors and of their catalytic properties is quite uneven, with some organelles such as the mitochondrion or the nucleus displaying higher occurrence than the others. The main localization of iron-dependent enzymes, which constitute 6.5% of all human enzymes, is the endoplasmic reticulum, where they catalyze the modification of both endo- and exogenous molecules and metabolites. Human iron-enzymes have a lower number of protein residues in their IBPs, in order to allow the iron ion to coordinate directly to the substrate.

Materials and methods

Proteins are generally composed of one or more functional regions, commonly termed domains. The identification of domains that occur within proteins can therefore provide insights into their function. Pfam is a database of protein domains, defined on the basis of the comparison of ensembles of protein regions that share a significant degree of sequence similarity, thereby suggesting homology. Each domain is represented by a multiple sequence alignment and by a more complex mathematical representation called a hidden Markov model (HMM). HMMs can be used for analyzing proteomes to search for occurrences of the corresponding domain (see below). Each domain entry in the Pfam database has an annotation, which may include the ability to bind metal cofactors.

Using the approach described in ref. 39 as implemented in the RDGB program,[40] we predicted all iron-binding proteins (IBPs) encoded by the human genome. RDGB is a computational tool written in Python. The approach of RDGB exploits the protein domains of the Pfam database to identify putative homologues of the proteins of interest in any desired genome or list of genomes. Thus, the input to RDGB is a list of Pfam domains of interest (in our case, domains associated with iron-binding capability) and a list of genomes to be analyzed (in our case only the human genome).

The input list of Pfam domains is created by merging two lists: first, the list of all Pfam domains annotated as iron-binding, retrieved by mining the text of the annotations in the database; second, from the analysis of the sequence of iron-binding proteins with known 3D structure that are available from the Protein Data Bank (PDB). In the latter case, we extract from the PDB database also the pattern of amino acids that are responsible for metal binding (*i.e.* the metal binding pattern, MBP) and its position within the domain sequence. The MBP is defined by the identity and spacing of the amino acids, *e.g.*, CX4CX20H, where X is any amino acid. This pattern provides a way to filter the initial results in order to reduce the number of false positives[39] (*i.e.*, of the proteins containing a Pfam domain annotated as iron-binding but which in reality are unable to bind it) by rejecting the proteins that lack the MBP or that have the MBP in the wrong position within the domain. The MBP filter cannot be applied in the absence of a relevant 3D structure available from the PDB. The MetalPDB database contains information on all the MBPs and the Pfam domains found in structurally characterized metalloproteins.[9] Our search started from 352 Pfam domains: 261 with an associated iron-containing 3D structure (102 binding individual iron ions, 80 binding iron–sulfur clusters, and 79 binding heme) and 91 annotated as iron-binding domains.

This search was integrated by locally searching from MBPs within all human protein sequences. This is done by extracting from the HMM representing the Pfam domain that contains the binding site of interest only the regions around the MBP. This "trimmed domain" provides a convenient way to search for a MBP regardless of the agreement with the whole Pfam domain, thus affording a better sensitivity in the detection of MBPs in divergent sequences.[41]

In total we retrieved 363 human iron-proteins. As a qualitative indicator of reliability of our dataset, we checked whether one of the following conditions applied (in decreasing order of reliability):

(1) A 3D structure of the human protein in the iron-bound form is available (105 proteins).

(2) A 3D structure of a close homolog (sequence identity $\geq$50%) of the human protein in the iron-bound form is available (76 proteins).

(3) The predicted protein contains an iron-binding Pfam domain with a conserved MBP (147 proteins).

(4) The predicted protein contains a conserved MBP (based on local search) (22 proteins).

(5) The predicted protein contains an iron-binding Pfam domain, but the occurrence of the MBP cannot be verified due to the lack of a 3D structure for that domain family (13 proteins).

We integrated these predictions by adding the proteins annotated in the Uniprot database, a public comprehensive resource of protein sequence and functional information, as "iron-binding", "iron–sulfur-binding", or "heme-binding". This contributed 35 additional iron-proteins.

For each predicted iron-protein, we retrieved the following annotations from UniProt:[42] intracellular location, EC number, biological processes as reported in the Gene Ontology database,[43] involvement in diseases. Further annotation such as the cofactor role and type were manually added by inspecting the literature. We used the Swiss-Prot database (at February 2018 contained 20259 entries)[34] to compare the iron-protein dataset with all human proteins. For the latter dataset, annotations were retrieved from Uniprot in the same way as for the iron-protein dataset.

The 3D structural model of the HSPB1-associated protein 1 was built using MODELER v.9.2[44] and energy-refined using the AMBER[45] web server provided by the WeNMR platform.[46]

Abbreviations

IBP	Iron-binding pattern
ROS	Reactive oxygen species
ROR	Retinoid-related orphan receptor

Conflicts of interest

There are no conflicts to declare.

References

1 C. Andreini, L. Banci, I. Bertini, S. Elmi and A. Rosato, Non-heme iron through the three domains of life, *Proteins: Struct., Funct., Bioinf.*, 2007, **67**, 317–324.
2 C. Andreini, A. Rosato and L. Banci, The relationship between environmental dioxygen and iron–sulfur proteins explored at the genome level, *PLoS One*, 2017, **12**, e0171279.
3 A. Camacho, X. A. Walter, A. Picazo and J. Zopfi, Photoferrotrophy: Remains of an Ancient Photosynthesis in Modern Environments, *Front. Microbiol.*, 2017, **8**, 323.
4 J. J. R. Frausto da Silva and R. J. P. Williams, *The Biological Chemistry of the Elements*, Oxford, Oxford, 1991.
5 I. Bertini, A. Sigel and H. Sigel, *Handbook on Metalloproteins*, New York, Marcel Dekker, 2001, p. 1800.
6 N. Maio and T. A. Rouault, Iron–sulfur cluster biogenesis in mammalian cells: new insights into the molecular mechanisms of cluster delivery, *Biochim. Biophys. Acta*, 2015, **1853**, 1493–1512.
7 T. A. Rouault, Mammalian iron-sulphur proteins: novel insights into biogenesis and function, *Nat. Rev. Mol. Cell Biol.*, 2015, **16**, 45–55.
8 S. Ciofi-Baffoni, V. Nasta and L. Banci, Protein networks in the maturation of human iron–sulfur proteins, *Metallomics*, 2018, **10**, 49–72.
9 V. Putignano, A. Rosato, L. Banci and C. Andreini, PDB in 2018: a database of metal sites in biological macromolecular structures, *Nucleic Acids Res.*, 2018, **46**, D459–D464.
10 S. J. Dixon and B. R. Stockwell, The role of iron and reactive oxygen species in cell death, *Nat. Chem. Biol.*, 2014, **10**, 9–17.
11 M. D. Knutson, Iron transport proteins: gateways of cellular and systemic iron homeostasis, *J. Biol. Chem.*, 2017, **292**, 12735–12743.
12 R. Coffey and T. Ganz, Iron homeostasis: an anthropocentric perspective, *J. Biol. Chem.*, 2017, **292**, 12727–12734.
13 H. Contreras, N. Chim, A. Credali and C. W. Goulding, Heme uptake in bacterial pathogens, *Curr. Opin. Chem. Biol.*, 2014, **19**, 34–41.
14 C. Ratledge and L. G. Dover, Iron metabolism in pathogenic bacteria, *Annu. Rev. Microbiol.*, 2000, **54**, 881–941.
15 T. A. Rouault and W. H. Tong, Iron–sulfur cluster biogenesis and human disease, *Trends Genet.*, 2008, **24**, 398–407.
16 M. Caza and J. W. Kronstad, Shared and distinct mechanisms of iron acquisition by bacterial and fungal pathogens of humans, *Front. Cell. Infect. Microbiol.*, 2013, **3**, 80.
17 C. Andreini, L. Banci and A. Rosato, Exploiting bacterial operons to illuminate human iron–sulfur proteins, *J. Proteome Res.*, 2016, **15**, 1308–1322.
18 C. Andreini and I. Bertini, A bioinformatics view of zinc enzymes, *J. Inorg. Biochem.*, 2012, **111**, 150–156.
19 R. G. Kranz, R. Lill, B. Goldman, G. Bonnard and S. Merchant, Molecular mechanisms of cytochrome *c* biogenesis: three distinct systems, *Mol. Microbiol.*, 1998, **29**, 383–396.
20 F. Fontanesi, I. C. Soto, D. Horn and A. Barrientos, Assembly of mitochondrial cytochrome *c*-oxidase, a complicated and highly regulated cellular process, *Am. J. Physiol.: Cell Physiol.*, 2006, **291**, C1129–C1147.
21 F. Fontanesi, I. C. Soto and A. Barrientos, Cytochrome *c* oxidase biogenesis: new levels of regulation, *IUBMB Life*, 2008, **60**, 557–568.
22 R. Lill, R. Dutkiewicz, S. A. Freibert, T. Heidenreich, J. Mascarenhas, D. J. Netz, V. D. Paul, A. J. Pierik, N. Richter, M. Stumpfig, V. Srinivasan, O. Stehling and U. Muhlenhoff, The role of mitochondria and the CIA machinery in the maturation of cytosolic and nuclear iron–sulfur proteins, *Eur. J. Cell Biol.*, 2015, **94**, 280–291.
23 O. Stehling, C. Wilbrecht and R. Lill, Mitochondrial iron–sulfur protein biogenesis and human disease, *Biochimie*, 2014, **100**, 61–77.
24 O. Stehling and R. Lill, The role of mitochondria in cellular iron–sulfur protein biogenesis: mechanisms, connected processes, and diseases, *Cold Spring Harbor Perspect. Biol.*, 2013, **5**, a011312.
25 F. Camponeschi, S. Ciofi-Baffoni and L. Banci, Anamorsin/Ndor1 Complex Reduces [2Fe–2S]-MitoNEET via a Transient

Protein–Protein Interaction, *J. Am. Chem. Soc.*, 2017, **139**, 9479–9482.

26 V. Nasta, A. Giachetti, S. Ciofi-Baffoni and L. Banci, Structural insights into the molecular function of human (2Fe–2S) BOLA1–GRX5 and (2Fe–2S) BOLA3–GRX5 complexes, *Biochim. Biophys. Acta*, 2017, **1861**, 2119–2131.

27 M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin and G. Sherlock, Gene ontology: tool for the unification of biology. The Gene Ontology Consortium, *Nat. Genet.*, 2000, **25**, 25–29.

28 The Gene Ontology Consortium, The Gene Ontology project in 2008, *Nucleic Acids Res.*, 2008, **36**, D440–D444.

29 L. Banci, S. Ciofi-Baffoni, K. Gajda, R. Muzzioli, R. Peruzzini and J. Winkelmann, N-terminal domains mediate [2Fe–2S] cluster transfer from glutaredoxin-3 to anamorsin, *Nat. Chem. Biol.*, 2015, **11**, 772–778.

30 K. I. Pardee, X. Xu, J. Reinking, A. Schuetz, A. Dong, S. Liu, R. Zhang, J. Tiefenbach, G. Lajoie, A. N. Plotnikov, A. Botchkarev, H. M. Krause and A. Edwards, The structural basis of gas-responsive transcription by the human nuclear hormone receptor REV-ERBbeta, *PLoS Biol.*, 2009, **7**, e43.

31 J. A. Kallen, J. M. Schlaeppi, F. Bitsch, S. Geisse, M. Geiser, I. Delhon and B. Fournier, X-ray structure of the hRORalpha LBD at 1.63 A: structural and functional data that cholesterol or a cholesterol derivative is the natural ligand of RORalpha, *Structure*, 2002, **10**, 1697–1707.

32 C. Stehlin-Gaon, D. Willmann, D. Zeyer, S. Sanglier, D. A. Van, J. P. Renaud, D. Moras and R. Schule, All-trans retinoic acid is a ligand for the orphan nuclear receptor ROR beta, *Nat. Struct. Biol.*, 2003, **10**, 820–825.

33 M. B. van Niel, B. P. Fauber, M. Cartwright, S. Gaines, J. C. Killen, O. Rene, S. I. Ward, B. G. de Leon, Y. Deng, C. Eidenschenk, C. Everett, E. Gancia, A. Ganguli, A. Gobbi, J. Hawkins, A. R. Johnson, J. R. Kiefer, H. La, P. Lockey, M. Norman, W. Ouyang, A. Qin, N. Wakes, B. Waszkowycz and H. Wong, A reversed sulfonamide series of selective RORc inverse agonists, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 5769–5776.

34 A. Bairoch and R. Apweiler, The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000, *Nucleic Acids Res.*, 2000, **28**, 45–48.

35 R. Davydov, S. Im, M. Shanmugam, W. A. Gunderson, N. M. Pearl, B. M. Hoffman and L. Waskell, Role of the Proximal Cysteine Hydrogen Bonding Interaction in Cytochrome P450 2B4 Studied by Cryoreduction, Electron Paramagnetic Resonance, and Electron-Nuclear Double Resonance Spectroscopy, *Biochemistry*, 2016, **55**, 869–883.

36 J. H. Dawson and K. S. Eble, Cytochrome P450: heme iron coordination structure and mechanism of action, in *Advances in Inorganic and Bioinorganic Mechanism*, ed. Sykes J., London, Academic Press, 1986, pp. 2–64.

37 H. Li and T. L. Poulos, Structural variation in heme enzymes: a comparative analysis of peroxidase and P450 crystal structures, *Structure*, 1994, **2**, 461–464.

38 R. Lill and U. Muhlenhoff, Iron–sulfur protein biogenesis in eukaryotes: components and mechanisms, *Annu. Rev. Cell Dev. Biol.*, 2006, **22**, 457–486.

39 C. Andreini, I. Bertini and A. Rosato, Metalloproteomes: a bioinformatic approach, *Acc. Chem. Res.*, 2009, **42**, 1471–1479.

40 C. Andreini, I. Bertini, G. Cavallaro, L. Decaria and A. Rosato, A simple protocol for the comparative analysis of the structure and occurrence of biochemical pathways across superkingdoms, *J. Chem. Inf. Model.*, 2011, **51**, 730–738.

41 Y. Valasatava, A. Rosato, L. Banci and C. Andreini, Metalpredator: a web server to predict iron–sulfur cluster binding proteomes, *Bioinformatics*, 2016, btw 238.

42 The Uniprot Consortium, UniProt: the universal protein knowledgebase, *Nucleic Acids Res.*, 2017, **45**, D158–D169.

43 E. Camon, M. Magrane, D. Barrell, V. Lee, E. Dimmer, J. Maslen, D. Binns, N. Harte, R. Lopez and R. Apweiler, The Gene Ontology Annotation (GOA) Database: sharing knowledge in Uniprot with Gene Ontology, *Nucleic Acids Res.*, 2004, **32**, D262–D266.

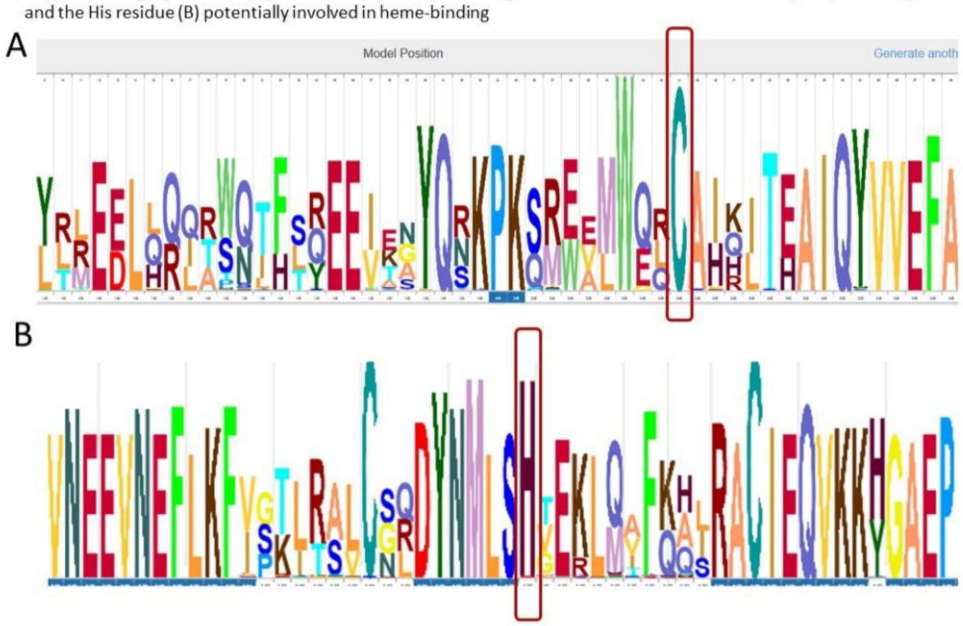
44 B. Webb and A. Sali, Comparative Protein Structure Modeling Using MODELLER, *Curr. Protoc. Bioinformatics*, 2014, **47**, 5.

45 D. A. Case, T. A. Darden, T. E. Cheatham, III, C. L. Simmerling and J. Wang, *et al.*, *AMBER 12, version 12*, University of California, San Francisco, CA, 2012.

46 I. Bertini, D. A. Case, L. Ferella, A. Giachetti and A. Rosato, A grid-enable web portal for NMR structure refinement with AMBER, *Bioinformatics*, 2011, **27**, 2384–2390.

Article supplementary material

Supplementary Figure S1: Skylign of the multiple sequence alignment of the mammalian ROR family only those segments including the CYS residue (A) and the His residue (B) potentially involved in heme-binding



Captions to Supplementary Tables

Table S1: List of all human proteins binding individual iron ions. Column 1 is a sequential number; column 2 reports the identifier (*Uniprot ID*) of the protein in the Uniprot (https://www.uniprot.org/) database; column 3 (*Confidence level*) summarizes the evidence supporting the assignment of the protein as an iron-protein, which is detailed in the next columns; columns 4 and 5 report for proteins that have been structurally characterized or that have a structurally characterized homolog in the Protein Data Bank (PDB) the PDB identifier and the percentage of sequence identity between the human protein and that homolog (only structures containing iron have been taken into account); columns 6 to 8 (*Method 2*, *3* and *4*, respectively) report the search results returned by each method. Columns 6 to 8 have been populated only for proteins that do not have entries in columns 4 and 5. Column 6 refers to the results of Pfam domain searches, after filtering for a known iron-binding pattern (IBP); the name of the Pfam domain and the location of the IBP within the sequence of the predicted human iron-protein are reported. Column 7 refers to the results of local sequence searches, based on the occurrence of a known iron-binding pattern (IBP); the location of the IBP within the sequence of the predicted human iron-protein is reported. Column 8 refers to the results of

Pfam domain searches, for domains lacking an associated iron-binding pattern (IBP); the name of the Pfam domain is reported.

Table S2: List of all human heme-binding proteins. For details see the caption to Supplementary Table S1.

Table S3: List of all human iron-sulfur proteins. For details see the caption to Supplementary Table S1.

Table S4: Functional properties of the human proteins binding individual iron ions. Column is a sequential number; column 2 reports the identifier (*Uniprot ID*) of the protein in the Uniprot (https://www.uniprot.org/) database; column 3 (*Entry name*) reports the name of this entry in Uniprot; column 4 (*Gene names*) reports the name of the gene, together with all its alternative names in Uniprot, coding for the protein; column 5 (*Protein name*) reports the name of the protein, together with all its alternative names in Uniprot; column 6 (*Predicted pattern*) reports the predicted IBP; column 7 reports the number of iron ions predicted to be in the physiological metal site(s); column 8 (*Iron role*) reports the physiological role of the iron site; column 9 (*EC number*) reports the Enzyme Commission number for iron-dependent enzymes; column 10 reports the subcellular location(s) of the protein; column 11 specifies whether the protein is associated to the membrane; column 12 (*Involvement in disease*) reports the disease annotation in Uniprot; column 13 (*Gene ontology*) reports the terms from the Gene Ontology database associated to the biological processes involving the protein.

Table S5: Functional properties of the human heme-binding proteins. For details see the caption to Supplementary Table S4.

Table S6: Functional properties of the human iron-sulfur proteins. For details see the caption to Supplementary Table S4.

Table S1: *List of all human proteins binding individual iron ions.*

					Prediction methods are reported from the most reliable to the less reliable (from left to right)		
			Method 1		Method 2	Method 3	Method 4
	Uniprot ID	Confidence level	Fe-binding pdb_chain	Sequence Id with a Fe-binding pdb_chain	Contains a Fe-binding domain with conserved ligands level	Contains a known iron-binding site	Contains a Fe-binding domain with unknown ligands
1	PHYD1_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3obz_A	100			
2	PIR_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1jli_A	100			
3	UTY_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3zli_A	100			
4	KDM6B_HUMAN	A 3D structure of the human protein in the iron-bound form is available	2xue_A	100			
5	KDM4A_HUMAN	A 3D structure of the human protein in the iron-bound form is available	5anq_A	100			
6	KDM4C_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4xdo_A	100			
7	KDM7A_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3kv5_A	100			
8	JMJD6_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3id8_A	100			
9	PHF2_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3pu8_A	100			
10	PAHX_HUMAN	A 3D structure of the human protein in the iron-bound form is available	2a1x_A	100			
11	PHF8_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3kv4_A	100			
12	EGLN1_HUMAN	A 3D structure of the human protein in the iron-bound form is available	2y34_A	100			
13	HIF1N_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1h2k_A	100			
14	TPH2_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4v06_A	100			
15	TPH1_HUMAN	A 3D structure of the human protein in the iron-bound form is available	5j6d_A	100			
16	DOHH_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4d4z_A	100			
17	GSTP1_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1zgn_A	100			
18	FRIH_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4oyn_A	100			
19	TRFL_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1bka_A	100			
20	LX15B_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4nre_A	100			
21	MTND_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4qgn_A	100			
22	RIR2B_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3hf1_A	100			
23	PP2BA_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1aui_A	100			
24	PP2BB_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4or9_A	100			
25	RPE_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3ovp_A	100			
26	FBXL5_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3v5x_A	100			
27	TET2_HUMAN	A 3D structure of the human protein in the iron-bound form is available	5d9y_A	100			
28	LOX12_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3d3l_A	99			
29	FTO_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3lfm_A	99			
30	KDM4D_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3dxu_A	99			
31	KDM2A_HUMAN	A 3D structure of the human protein in the iron-bound form is available	2yu1_A	99			
32	TRFE_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3v83_A	99			
33	HGD_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1ey2_A	99			
34	PPA5_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1war_A	99			
35	HEMH_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3w1w_A	99			
36	RIOX1_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4e4h_A	99			
37	Q7KZA3_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3w1w_A	99			
38	ETHE1_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4chl_A	98			
39	LOX5_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3o8y_A	97			
40	ALKB3_HUMAN	A 3D structure of the human protein in the iron-bound form is available	2iuw_A	97			
41	RPEL1_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3ovp_A	96			
42	KDM6A_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4ufO_A	94			
43	RPE65_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	3fsn_A	98			
44	PH4H_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	5den_A	92			
45	TY3H_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1toh_A	91			
46	RIR2_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1w68_A	91			
47	HPPD_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1sqi_A	89			
48	MIOX_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	2huo_A	89			
49	3HAO_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	3fe5_A	86			
50	KDM4E_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	3dxu_A	84			
51	KDM4B_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4xdo_A	83			
52	LOX15_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	2p0m_A	81			
53	PP2BC_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1aui_A	81			
54	FTMT_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4oyn_A	80			
55	TET3_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	5d9y_A	72			
56	TET1_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	5d9y_A	68			
57	KDM2B_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	2yu1_A	66			
58	FHL19_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4oyn_A	66			
59	FHL17_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4oyn_A	65			
60	EGLN3_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	2g19_A	64			
61	EGLN2_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	2y34_A	64			
62	FRIL_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4mjy_A	60			
63	KDM5A_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4igo_A	56			
64	GALT_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1hxq_A	56			
65	D3DRM8_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1hxq_A	56			
66	KDM5B_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4igo_A	55			
67	KDM5C_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4igo_A	54			
68	KDM5D_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4igo_A	54			
69	MAP11_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	3s6b_A	53			
70	LOXE3_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4nre_A	51			
71	TRFM_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Transferrin (D78-Y107-Y210-H279), Transferrin (Y451-Y556-H625)		
72	TMLH_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			TauD (H242-D244-H389)		
73	BODG_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			TauD (H202-D204-H347)		
74	BCDO2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			RPE65 (H226-H286-H357-H573)		
75	BCDO1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			RPE65 (H172-H237-H308-H514)		
76	MAP2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Peptidase_M24 (D251-D262-H331-E364-E459)		

77	MAP12_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Peptidase_M24 (D178-D189-H252-E284-E315)		
78	OSGEP_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Peptidase_M22 (H109-H113-Y130-D294)		
79	NIF3L_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			NIF3 (H93-H339-E343)		
80	K1456_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Methyltransf_11 (H112)		
81	MRE11_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Metallophos_2 (D20-H22-D60)		
82	MPPD1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Metallophos (D97-H99-D118-H286)		
83	PP1A_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Metallophos (D64-H66-D92)		
84	TMM62_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Metallophos (D63-H65-D99)		
85	TMPPE_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Metallophos (D214-H216-D246-H393)		
86	ACP7_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Metallophos (D141-D170-Y173-H335)		
87	LX12B_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Lipoxygenase (H398-H403-H578)		
88	KDM3B_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			JmjC (H1604-H1689)		
89	JMJD4_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			JmjC (H235-D237-H315)		
90	HOT_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Fe-ADH (D242-H246-H330-H357)		
91	KDM8_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Cupin_8 (H321-D323-H400)		
92	JMJD8_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Cupin_8 (H249-H251-H318)		
93	JMJD7_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Cupin_8 (H178-D180-H277)		
94	HBAP1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			JmjC (H175-D177-H257)		
95	TYW5_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Cupin_8 (H160-D162-H235)		
96	HUTI_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Amidohydro_3 (H87-H89), Amidohydro_3 (H260-H283-D334), Amidohydro_1 (H87-H89-H260-D334)		
97	P3H1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy_3 (H587-D589-H659)		
98	P3H3_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy_3 (H584-D586-H656)		
99	P3H2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy_3 (H580-D582-H652)		
100	P4HA3_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy_3 (H440-D442-H510)		
101	P4HA2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy_3 (H430-D432-H501)		
102	P4HA1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy_3 (H429-D431-H500)		
103	P4HTM_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy_3 (H328-D330-H441)		
104	OGFD3_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy_3 (H230-D232-H288)		
105	OGFD1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy_3 (H155-D157-H218)		

106	ALKB8_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy_2 (H238-D240-H292)		
107	ALKB1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy_2 (H231-D233-H287)		
108	ALKB5_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy_2 (H204-D206-H266)		
109	ALKB2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy_2 (H171-D173-H236)		
110	ALKB4_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy_2 (H169-D171-H254)		
111	ALKB7_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy_2 (H121-D123-H177)		
112	ALKB6_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy_2 (H114-D116-H182)		
113	PLOD3_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy (H667-D669-H719)		
114	PLOD2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy (H666-D668-H718)		
115	PLOD1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2OG-FeII_Oxy (H656-D658-H708)		
116	JHD2C_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			JmjC (H2336-E2338-H2466)		
117	RIOX2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			JmjC (H179-D181-H240)		
118	KDM3A_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			JmjC (H1120-D1122-H1249)		
119	HAIR_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			JmjC (C1007-E1009-H1125)		
120	COQ7_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			COQ7 (E60-E90-H93-E142-E178-H181)		
121	ASPH_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Asp_Arg_Hydrox (H679-H725)		
122	ASPH2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Asp_Arg_Hydrox (H283-H328)		
123	NGAL_HUMAN	The predicted protein contains a conserved MBP (based on local search)				Y126-K145-K154	
124	SCD5_HUMAN	The predicted protein contains a conserved MBP (based on local search)				H94-H99-H131-H134-H135-H243-H272-H275-H276	
125	OGFD2_HUMAN	The predicted protein contains a conserved MBP (based on local search)				H235-D237-H290	
126	CH25H_HUMAN	The predicted protein contains a conserved MBP (based on local search)				H143-H147-H157-H161-H205-H238-H242-H243	
127	SC5D_HUMAN	The predicted protein contains a conserved MBP (based on local search)				H138-H142-H151-H155-H209-H228-H232-H233	
128	ACOD_HUMAN	The predicted protein contains a conserved MBP (based on local search)				H120-H125-H157-H160-H161-H269-H298-H301-H302	
129	AEDO_HUMAN	The predicted protein contains a conserved MBP (based on local search)				H112-H114-H193	
130	HPDL_HUMAN	The predicted protein contains a conserved MBP (based on local search)				H163-H258-E339	
131	NRAM2_HUMAN	The predicted protein contains an iron-binding Pfam domain, but the occurrence of the MBP cannot be verified due to the lack of a 3D structure for that domain family					Nramp
132	NRAM1_HUMAN	The predicted protein contains an iron-binding Pfam domain, but the occurrence of the MBP cannot be verified due to the lack of a 3D structure for that domain family					Nramp
133	MSMO1_HUMAN	Annotated as iron-binding in Uniprot (pubmed id 20643956)					
134	ALKMO_HUMAN	Annotated as iron-binding in Uniprot (pubmed id 8663358)					
135	FRDA_HUMAN	Annotated as iron-binding in Uniprot (pubmed id 15641778)					
136	S40A1_HUMAN	Annotated as iron-binding in Uniprot (pubmed id 12091367)					
137	HEPC_HUMAN	Annotated as iron-binding in Uniprot (pubmed id 16009582)					
138	MFRN2_HUMAN	Annotated as iron-binding in Uniprot					
139	MFRN1_HUMAN	Annotated as iron-binding in Uniprot					

Table S2: *List of all human heme-binding proteins.*

					Prediction methods are reported from the most reliable to the less reliable (from left to right)		
			Method 1		Method 2	Method 3	Method 4
	Uniprot ID	Confidence level	Fe-binding pdb_chain	Sequence Id with a Fe-binding pdb_chain	Contains a Fe-binding domain with conserved ligands level	Contains a known iron-binding site	Contains a Fe-binding domain with unknown ligands
1	CATA_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1f4j_A	100			
2	CP17A_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3ruk_A	100			
3	CP11A_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3n9y_A	100			
4	CP19A_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3eqm_A	100			
5	PTGIS_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3b6h_A	100			
6	NOS3_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4d1o_A	100			
7	CP2R1_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3c6g_A	100			
8	CP46A_HUMAN	A 3D structure of the human protein in the iron-bound form is available	2q9f_A	100			
9	CP2D6_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3qm4_A	100			
10	CP7A1_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3dax_A	100			
11	CP1A1_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4i8v_A	100			
12	CP1A2_HUMAN	A 3D structure of the human protein in the iron-bound form is available	2hi4_A	100			
13	CP51A_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3jus_A	100			
14	NOS2_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1nsi_A	100			
15	PERM_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3f9p_C	100			
16	PGRC1_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4x8y_A	100			
17	CYB5B_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3ner_A	100			
18	CYC_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1j3s_A	100			
19	HBG2_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1fdh_G	100			
20	CYGB_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3ag0_A	100			
21	NGB_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4mpm_A	100			
22	HBG1_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1i3d_A	100			
23	HBAZ_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3w4u_A	100			
24	HBD_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1shr_B	100			
25	THAP4_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3ia8_A	100			
26	ALBU_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1n5u_A	100			
27	CBS_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4l3v_A	100			
28	CBSL_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4l3v_A	100			
29	I23O1_HUMAN	A 3D structure of the human protein in the iron-bound form is available	5ek2_A	100			
30	Q6LEN0_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3b6h_A	100			
31	Q5HYD9_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3ner_A	100			
32	HEMH_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3w1w_A	99			
33	CP21A_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4y8w_A	99			
34	C11B2_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4zgx_A	99			
35	NOS1_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4uh5_A	99			
36	CP2C9_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1r9o_A	99			
37	CP2CJ_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4gqs_A	99			
38	CP3A4_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3tjs_A	99			
39	HMOX1_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4wd4_C	99			
40	HMOX2_HUMAN	A 3D structure of the human protein in the iron-bound form is available	2qpp_A	99			
41	NR1D2_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3cqv_A	99			
42	CP2E1_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3e4e_A	98			
43	CP1B1_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3pm0_A	98			
44	CP2C8_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1pq2_A	98			
45	CP2B6_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3ibd_A	98			
46	CP2AD_HUMAN	A 3D structure of the human protein in the iron-bound form is available	2p85_A	98			
47	CP2A6_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1z10_A	98			
48	NB5R4_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3lf5_A	98			
49	MYG_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3rgk_A	98			
50	Q14412_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4mqj_B	97			
51	Q13120_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1z10_A	95			
52	HBB_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1dxt_B	100			
53	HBA_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1bz1_A	100			
54	HBE_HUMAN	A 3D structure of the human protein in the iron-bound form is available	1a9w_E	100			
55	C11B1_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4zgx_A	92			
56	CP2A7_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1z10_A	92			
57	CY1_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4d6u_D	92			
58	PGH1_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1cqe_A	92			
59	COX1_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1occ_A	91			
60	C560_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4ytp_C	91			
61	Q14097_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	3ibd_A	91			
62	CP2D7_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	3qm4_A	90			
63	CYB5_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	2m33_A	90			
64	CP3A7_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	3tjs_A	88			
65	DHSD_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4ytp_D	88			
66	PGH2_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1pxx_A	86			
67	Q7Z2Y6_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	5b72_A	85			
68	CP24A_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	3k9v_A	84			
69	PERL_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	2ikc_A	84			
70	CP3A5_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	3tjs_A	83			
71	CP2CI_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1r9o_A	81			
72	HEMO_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1qhu_A	79			
73	CYB_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1be3_C	78			

74	Q16750_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4gqs_A	77			
75	CP343_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	3tjs_A	75			
76	PERE_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1cxp_C	72			
77	NR1D1_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	3cqv_A	71			
78	SUOX_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1sox_A	68			
79	PGRC2_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4x8y_A	68			
80	Q7Z348_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1dt6_A	66			
81	HBAT_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	3fh9_A	65			
82	C2G1P_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4h1n_A	62			
83	T23O_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4hka_A	59			
84	CP2F1_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	2p85_A	53			
85	HBM_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1v75_A	52			
86	CP2S1_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	2q6n_A	50			
87	CP2J2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (C448)		
88	NEUFC_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Cyt-b5 (Y79)		
89	CP2U1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (C490)		
90	FETA_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Serum_albumin (Y185-Y377)		
91	CP8B1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (H120-C440)		
92	CP2W1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (C433)		
93	CYAC3_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Cytochrom_B561 (H47-H83-H117-H156)		
94	GCYB2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			HNOB (H26)		
95	FS2P1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Cyt-b5 (H90-H113)		
96	NENF_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Cyt-b5 (Y88)		
97	THAS_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (C479)		
98	CYBR1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Cytochrom_B561 (H50-H86-H120-H159)		

99	CP26C_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (H138-C459)		
100	CP26B_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (H138-C441)		
101	GCYB1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			HNOB (H105)		
102	CP27B_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (C455)		
103	CP26A_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (H133-C442)		
104	CY561_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Cytochrom_B561 (H53-H87-H121-H160)		
105	CP27A_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (C476)		
106	FADS3_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Cyt-b5 (H55-H78-H186)		
107	C27C1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (C318)		
108	CP4Z1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (C452)		
109	CP4FN_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (E335-C475)		
110	CP4F8_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (C468)		
111	CP4FC_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (C468)		
112	CP4AB_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (E321-C457)		
113	CP4F2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (E328-C468)		
114	CP4AM_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (E321-C457)		
115	CP4F3_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (E328-C468)		
116	CP4FB_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (E328-C468)		
117	CP4V2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (E329-C467)		
118	CP4X1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (C454)		
119	CP4B1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (E315-C453)		
120	PERT_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			An_peroxidase (H494)		
121	PXDN_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			An_peroxidase (H1074)		
122	CB5D1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Cyt-b5 (Y52-H83)		
123	CS6D1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Cytochrom_B561 (H55-H93-H127-H166)		

124	C56D2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Cytochrom_B561 (H48-H86-H120-H159)		
125	PER1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			PAS (H409)		
126	I23O2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			IDO (H360)		
127	RORG_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Hormone_recep (H479)		
128	RORA_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Hormone_recep (H484)		
129	CP7B1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (C449)		
130	RORB_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Hormone_recep (H434)		
131	PXDNL_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			An_peroxidase (H1057)		
132	CY24A_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Cytochrom_B558a (H94)		
133	AFAM_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Serum_albumin (Y377)		
134	CP39A_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (C414)		
135	CP20A_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			p450 (C409)		
136	FRRS1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Cytochrom_B561 (H373-H414-H446-H482)		
137	CP052_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Cyt_bd_oxida_I (E125)		
138	FRS1L_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			DOMON (M205)		
139	MOXD1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			DOMON (M70)		
140	DUOX2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Ferric_reduct (H774-H1222-H1235)		
141	DUOX1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Ferric_reduct (H770-H1225-H1238)		
142	STEA3_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Ferric_reduct (H316-H409)		
143	CY24B_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Ferric_reduct (H101-H115-H209-H222)		
144	NOX1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Ferric_reduct (H101-H115-H209-H221)		
145	STEA2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Ferric_reduct (H316-H409)		
146	NOX4_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Ferric_reduct (H105-H119-H194-H207)		
147	STEA1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Ferric_reduct (H175-H268)		
148	STEA4_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Ferric_reduct (H304-H397)		

149	NOX5_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Ferric_reduct (H314-H328-H402-H415)		
150	NPAS2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			PAS_3 (H119-H171)		
151	FADS1_HUMAN	The predicted protein contains a conserved MBP (based on local search)				H52-H75-H138-H183	
152	FADS2_HUMAN	The predicted protein contains a conserved MBP (based on local search)				H53-H76-H184	
153	FA2H_HUMAN	The predicted protein contains a conserved MBP (based on local search)				H43-H69	
154	SHIP2_HUMAN	The predicted protein contains a conserved MBP (based on local search)				C405	
155	GCYA2_HUMAN	The predicted protein contains a conserved MBP (based on local search)				H480	
156	FRAS1_HUMAN	The predicted protein contains a conserved MBP (based on local search)				H1799-H1945-H2080-H3301	
157	DGCR8_HUMAN	The predicted protein contains a conserved MBP (based on local search)				C352	
158	COX15_HUMAN	The predicted protein contains an iron-binding Pfam domain, but the occurrence of the MBP cannot be verified due to the lack of a 3D structure for that domain family					COX15-CtaA
159	COX5A_HUMAN	The predicted protein contains an iron-binding Pfam domain, but the occurrence of the MBP cannot be verified due to the lack of a 3D structure for that domain family					COX5A
160	CCHL_HUMAN	The predicted protein contains an iron-binding Pfam domain, but the occurrence of the MBP cannot be verified due to the lack of a 3D structure for that domain family					Cyto_heme_lyase
161	Q68D50_HUMAN	The predicted protein contains an iron-binding Pfam domain, but the occurrence of the MBP cannot be verified due to the lack of a 3D structure for that domain family					Cyto_heme_lyase
162	GCYA3_HUMAN	The predicted protein contains an iron-binding Pfam domain, but the occurrence of the MBP cannot be verified due to the lack of a 3D structure for that domain family					HNOB
163	HRG1_HUMAN	The predicted protein contains an iron-binding Pfam domain, but the occurrence of the MBP cannot be verified due to the lack of a 3D structure for that domain family					HRG
164	HEBP1_HUMAN	The predicted protein contains an iron-binding Pfam domain, but the occurrence of the MBP cannot be verified due to the lack of a 3D structure for that domain family					SOUL
165	HEBP2_HUMAN	The predicted protein contains an iron-binding Pfam domain, but the occurrence of the MBP cannot be verified due to the lack of a 3D structure for that domain family					SOUL
166	HRG_HUMAN	Annotated as heme-binding in Uniprot (pubmed id 678554)					
167	STC2_HUMAN	Annotated as heme-binding in Uniprot (pubmed id 22503972)					
168	BACH1_HUMAN	Annotated as heme-binding in Uniprot (pubmed id 21555518)					
169	SRC_HUMAN	Annotated as heme-binding in Uniprot (pubmed id 21036157)					
170	JAK2_HUMAN	Annotated as heme-binding in Uniprot (pubmed id 21036157)					
171	FLVC1_HUMAN	Annotated as heme-binding in Uniprot (pubmed id 20610401)					
172	FLVC2_HUMAN	Annotated as heme-binding in Uniprot (pubmed id 20610401)					
173	AMBP_HUMAN	Annotated as heme-binding in Uniprot (pubmed id 11877257)					
174	ABCB7_HUMAN	Annotated as heme-binding in Uniprot					
175	ABCB6_HUMAN	Annotated as heme-binding in Uniprot					
176	COPA_HUMAN	Annotated as heme-binding in Uniprot					
177	EMAL6_HUMAN	Annotated as heme-binding in Uniprot					
178	ADGB_HUMAN	Annotated as heme-binding in Uniprot					
179	C163A_HUMAN	Annotated as heme-binding in Uniprot					

180	ABCG2_HUMAN	Annotated as heme-binding in Uniprot					
181	PCFT_HUMAN	Annotated as heme-binding in Uniprot					
182	E2AK1_HUMAN	Annotated as heme-binding in Uniprot					
183	PGES2_HUMAN	Annotated as heme-binding in Uniprot					
184	KLKB1_HUMAN	Annotated as heme-binding in Uniprot					
185	HERC2_HUMAN	Annotated as heme-binding in Uniprot					
186	Q6ZNJ6_HUMAN	Annotated as heme-binding in Uniprot					
187	Q68D05_HUMAN	Annotated as heme-binding in Uniprot					
188	A0A024RAI7_HUMAN	Annotated as heme-binding in Uniprot					
189	Q658T6_HUMAN	Annotated as heme-binding in Uniprot					
190	Q8N3P5_HUMAN	Annotated as heme-binding in Uniprot					
191	CP4Z2_HUMAN	Annotated as heme-binding in Uniprot					
192	PER3_HUMAN	Annotated as heme-binding in Uniprot					

Table S3: *List of all human iron-sulfur proteins.*

					Prediction methods are reported from the most reliable to the less reliable (from left to right)		
			Method 1		Method 2	Method 3	Method 4
	Uniprot ID	Confidence level	Fe-binding pdb_chain	Sequence Id with a Fe-binding pdb_chain	Contains a Fe-binding domain with conserved ligands level	Contains a known iron-binding site	Contains a Fe-binding domain with unknown ligands
1	GLRX5_HUMAN	A 3D structure of the human protein in the iron-bound form is available	2wul_A	100			
2	XDH_HUMAN	A 3D structure of the human protein in the iron-bound form is available	2ckj_A	100			
3	AOXA_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4uhw_A	100			
4	ADX_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3p1m_A	100			
5	CISD1_HUMAN	A 3D structure of the human protein in the iron-bound form is available	2qd0_A	100			
6	MUTYH_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3n5n_X	100			
7	PRI2_HUMAN	A 3D structure of the human protein in the iron-bound form is available	4rr2_D	100			
8	ACOC_HUMAN	A 3D structure of the human protein in the iron-bound form is available	2b3x_A	100			
9	FDX2_HUMAN	A 3D structure of the human protein in the iron-bound form is available	2y5c_A	99			
10	HEMH_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3w1w_A	99			
11	CISD2_HUMAN	A 3D structure of the human protein in the iron-bound form is available	3fnv_A	97			
12	ACON_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1b0j_A	96			
13	NDUS2_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	5gpn_Z	95			
14	GABT_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1ohv_A	95			
15	ETFD_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	2gmh_A	95			
16	SDHB_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4ytp_B	95			
17	GLRX2_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	2ht9_A	93			
18	DPYD_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	1gt8_A	92			
19	UCRI_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4d6u_R	90			
20	UCRIL_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4d6u_R	89			
21	RFESD_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	3d89_A	88			
22	DNA2_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	5eaw_A	80			
23	ABCE1_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	3j16_B	68			
24	NDUS7_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	2fug_6	55			
25	IREB2_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	2b3x_A	53			
26	CISD3_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	3tbn_A	52			
27	ISCU_HUMAN	A 3D structure of a close homolog (sequence identity ≥ 50%) of the human protein in the iron-bound form is available	4eb5_C	50			
28	CDKAL_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			UPF0004 (C73-C109-C138); Radical_SAM (C214-C218-C221)		
29	CK5P1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			UPF0004 (C109-C145-C183); Radical_SAM (C258-C262-C265)		
30	AIFM3_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Rieske (C109-H111-C128-H131)		
31	RSAD2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Radical_SAM (C83-C87-C90)		
32	MOCS1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Radical_SAM (C80-C84-C87); Mob_synth_C (C312-C315-C329)		
33	RSAD1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Radical_SAM (C49-C53-C56)		
34	TYW1B_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Radical_SAM (C352-C356-C359)		
35	LIAS_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			LIAS_N(C106-C111-C117), Radical_SAM (C137-C141-C144)		
36	ELP3_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Radical_SAM (C99-C109-C112)		
37	NDUV1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			NADH_4Fe-4S (C379-C382-C385-C425)		
38	GLRX3_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Glutaredoxin (C159; C261)		
39	NDUS8_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Fer4_7 (C121-C150-C153-C156; C111-C114-C117-C160)		
40	NDUS1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Fer2_4 (C64-C75-C78-C92); NADH-G_4Fe-4S_3(H124-C128-C131-C137)	C176-C179-C182-C226	
41	NARFL_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Fe_hyd_lg_C (C190-C246-C395-C399)	C24-C71-C74-C77	
42	NARF_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Fe_hyd_lg_C (C172-C228-C374-C378)		
43	FANCJ_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			DEAD_2 (C283-C298-C310-C350)		
44	RTEL1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			DEAD_2 (C145-C163-C172-C207)		
45	ERCC2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			DEAD_2 (C116-C134-C155-C190)		
46	NFS1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Aminotran_5 (C381)		
47	NDUV2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			2Fe-2S_thioredx (C135-C140-C176-C180)		
48	ISCA2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Fe-S_biosyn (C79-C144-C146)		
49	ISCA1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Fe-S_biosyn (C57-C121-C123)		
50	DPH1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Diphthamide_syn (C115-C219-C347)		
51	DDX12_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			DEAD_2 (C286-C304-C334-C369)		
52	DDX11_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			DEAD_2 (C267-C285-C315-C350)		
53	DPH2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Diphthamide_syn (C88-C341)		
54	DPOLA_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			zf-DNA_Pol (C1348-C1353-C1371-C1374)		
55	NUBP2_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			ParA (C196-C199)		
56	NUBP1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			ParA (C235-C238)	C8-C22-C25-C31	
57	NUBPL_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			ParA (C244-C247)		
58	GRCR1_HUMAN	The predicted protein contains an iron-binding Pfam domain with a conserved MBP			Glutaredoxin (C156)		
59	NTH_HUMAN	The predicted protein contains a conserved MBP (based on local search)				C290-C297-C300-C306	
60	PUR1_HUMAN	The predicted protein contains a conserved MBP (based on local search)				C280-C426-C503-C506	
61	CPIN1_HUMAN	The predicted protein contains a conserved MBP (based on local search)				C237-C246-C249-C251	
62	NFU1_HUMAN	The predicted protein contains a conserved MBP (based on local search)				C210-C213	
63	REV3L_HUMAN	The predicted protein contains a conserved MBP (based on local search)				C1348-C1353-C1371-C1374	
64	DPOD1_HUMAN	The predicted protein contains a conserved MBP (based on local search)				C1058-C1061-C1071-C1076	
65	DPOE1_HUMAN	The predicted protein contains a conserved MBP (based on local search)				C2221-C2224-C2236-C2238	
66	BOLA3_HUMAN	The predicted protein contains an iron-binding Pfam domain, but the occurrence of the MBP cannot be verified due to the lack of a 3D structure for that domain family					BolA
67	BOLA2_HUMAN	The predicted protein contains an iron-binding Pfam domain, but the occurrence of the MBP cannot be verified due to the lack of a 3D structure for that domain family					BolA
68	BOLA1_HUMAN	The predicted protein contains an iron-binding Pfam domain, but the occurrence of the MBP cannot be verified due to the lack of a 3D structure for that domain family					BolA
69	ABCB7_HUMAN	Annotated as ironsulfur-binding in Uniprot					
70	CMAH_HUMAN	Annotated as ironsulfur-binding in Uniprot					

Table S4: *Functional properties of the human proteins binding individual iron ions.*

	Uniprot Id	Entry name	Gene names	Protein names	Predicted Pattern	Number of iron ions	Iron role	EC number	Subcellular location	Membrane associated	Involvement in disease	Gene ontology (biological process)
1	P46952	3HAO_HUMAN	HAAO	3-hydroxyanthranilate 3,4-dioxygenase (EC 1.13.11.6) (3-hydroxyanthranilate oxygenase) (3-HAO) (3-hydroxyanthranilic acid dioxygenase) (HAD)	H47-E53-H91	1 Fe cation	Catalytic	1.13.11.6	Cytoplasm	No		NAD biosynthetic process [GO:0009435]; neuron cellular homeostasis [GO:0070050]; quinolinate biosynthetic process [GO:0019805]; response to cadmium ion [GO:0046686]; response to zinc ion [GO:0010043]; tryptophan catabolic process [GO:0006569]
2	O00767	ACOD_HUMAN	SCD	Acyl-CoA desaturase (EC 1.14.19.1) (Delta(9)-desaturase) (Delta-9 desaturase) (Fatty acid desaturase) (Stearoyl-CoA desaturase) (hSCD1)	H120-H125-H157-H161; H160-H269-H298-H302	2 Fe cations	Catalytic	1.14.19.1	Endoplasmic reticulum	Yes		long-chain fatty-acyl-CoA biosynthetic process [GO:0035338]; unsaturated fatty acid biosynthetic process [GO:0006636]
3	Q6ZNF0	ACP7_HUMAN	ACP7 PAPL PAPL1	Acid phosphatase type 7 (EC 3.1.3.2) (Purple acid phosphatase long form)	D141-D170-Y173-H335	1 Fe cation	Catalytic	3.1.3.2	Extracellular space	No		
4	Q96SZ5	AEDO_HUMAN	ADO C10orf22	2-aminoethanethiol dioxygenase (EC 1.13.11.19) (Cysteamine dioxygenase)	H112-H114-H193	1 Fe cation	Catalytic	1.13.11.19	Unknown	No		oxidation-reduction process [GO:0055114]; sulfur amino acid catabolic process [GO:0000098]
5	Q13686	ALKB1_HUMAN	ALKBH1 ABH ABH1 ALKBH	Nucleic acid dioxygenase ALKBH1 (EC 1.14.11.-) (Alkylated DNA repair protein alkB homolog 1) (Alpha-ketoglutarate-dependent dioxygenase ABH1) (DNA 6mA demethylase) (DNA N6-methyl adenine demethylase) (EC 1.14.11.-) (DNA lyase ABH1) (EC 4.2.99.18) (DNA oxidative demethylase ALKBH1) (EC 1.14.11.33) (tRNA N1-methyl adenine demethylase) (EC 1.14.11.-)	H231-D233-H287	1 Fe cation	Catalytic	1.14.11.-; 4.2.99.18; 1.14.11.33	Mitochondrion, Nucleus	No		developmental growth [GO:0048589]; DNA dealkylation involved in DNA repair [GO:0006307]; DNA demethylation [GO:0080111]; DNA repair [GO:0006281]; in utero embryonic development [GO:0001701]; negative regulation of neuron apoptotic process [GO:0043524]; neuron migration [GO:0001764]; neuron projection development [GO:0031175]; oxidative demethylation [GO:0070989]; oxidative single-stranded DNA demethylation [GO:0035552]; placenta development [GO:0001890]; regulation of mitochondrial translation [GO:0070129]; regulation of translational elongation [GO:0006448]; regulation of translational initiation [GO:0006446]; RNA repair [GO:0042245]; tRNA demethylation [GO:1990983]; tRNA wobble cytosine modification [GO:0002101]
6	Q6NS38	ALKB2_HUMAN	ALKBH2 ABH2	DNA oxidative demethylase ALKBH2 (EC 1.14.11.33) (Alkylated DNA repair protein alkB homolog 2) (Alpha-ketoglutarate-dependent dioxygenase alkB homolog 2) (Oxy DC1)	H171-D173-H236	1 Fe cation	Catalytic	1.14.11.33	Nucleus	No		DNA dealkylation involved in DNA repair [GO:0006307]; DNA demethylation [GO:0080111]; oxidative demethylation [GO:0070989]; oxidative DNA demethylation [GO:0035511]
7	Q96Q83	ALKB3_HUMAN	ALKBH3 ABH3 DEPC1	Alpha-ketoglutarate-dependent dioxygenase alkB homolog 3 (EC 1.14.11.54) (Alkylated DNA repair protein alkB homolog 3) (hABH3) (DEPC-1) (Prostate cancer antigen 1)	H191-D193-H257	1 Fe cation	Catalytic	1.14.11.54	Cytoplasm, Nucleus	No		cell proliferation [GO:0008283]; DNA dealkylation involved in DNA repair [GO:0006307]; DNA repair [GO:0006281]; oxidative single-stranded DNA demethylation [GO:0035552]; oxidative single-stranded RNA demethylation [GO:0035553]
8	Q9NXW9	ALKB4_HUMAN	ALKBH4 ABH4	Alpha-ketoglutarate-dependent dioxygenase alkB homolog 4 (EC 1.14.11.-) (Alkylated DNA repair protein alkB homolog 4)	H169-D171-H254	1 Fe cation	Catalytic	1.14.11.-	Cytoplasm, Nucleus	No		actomyosin structure organization [GO:0031032]; cleavage furrow ingression [GO:0036090]; protein demethylation [GO:0006482]; regulation of transcription, DNA-templated [GO:0006355]; transcription, DNA-templated [GO:0006351]
9	Q6P6C2	ALKB5_HUMAN	ALKBH5 ABH5 OFOXD1	RNA demethylase ALKBH5 (EC 1.14.11.-) (Alkylated DNA repair protein alkB homolog 5) (Alpha-ketoglutarate-dependent dioxygenase alkB homolog 5)	H204-D206-H266	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No		cell differentiation [GO:0030154]; DNA dealkylation involved in DNA repair [GO:0006307]; mRNA export from nucleus [GO:0006406]; mRNA processing [GO:0006397]; oxidative single-stranded RNA demethylation [GO:0035553]; response to hypoxia [GO:0001666]; spermatogenesis [GO:0007283]
10	Q3KRA9	ALKB6_HUMAN	ALKBH6 ABH6	Alpha-ketoglutarate-dependent dioxygenase alkB homolog 6 (EC 1.14.11.-) (Alkylated DNA repair protein alkB homolog 6)	H114-D116-H182	1 Fe cation	Catalytic	1.14.11.-	Cytoplasm, Nucleus	No		
11	Q9BT30	ALKB7_HUMAN	ALKBH7 ABH7 SPATA11 UNQ6002/PRO34564	Alpha-ketoglutarate-dependent dioxygenase alkB homolog 7, mitochondrial (EC 1.14.11.-) (Alkylated DNA repair protein alkB homolog 7) (Spermatogenesis cell proliferation-related protein) (Spermatogenesis-associated protein 11)	H121-D123-H177	1 Fe cation	Catalytic	1.14.11.-	Mitochondrion	No		cellular response to DNA damage stimulus [GO:0006974]; fatty acid metabolic process [GO:0006631]; regulation of lipid storage [GO:0010883]; regulation of mitochondrial membrane permeability involved in programmed necrotic cell death [GO:1902445]
12	Q96BT7	ALKB8_HUMAN	ALKBH8 ABH8	Alkylated DNA repair protein alkB homolog 8 (EC 1.14.11.-) (Probable alpha-ketoglutarate-dependent dioxygenase ABH8) (S-adenosyl-L-methionine-dependent tRNA methyltransferase ABH8) (tRNA (carboxymethyluridine(34)-5-O)-methyltransferase ABH8) (EC 2.1.1.229)	H238-D240-H292	1 Fe cation	Catalytic	1.14.11.-; 2.1.1.229	Cytoplasm, Nucleus	No		cellular response to DNA damage stimulus [GO:0006974]; oxidation-reduction process [GO:0055114]; tRNA methylation [GO:0030488]; tRNA wobble uridine modification [GO:0002098]
13	Q6ZNB7	ALKMO_HUMAN	AGMO TMEM195	Alkylglycerol monooxygenase (EC 1.14.16.5) (Transmembrane protein 195)	H157-H161-H170-H173; H228-H250-H252-H253	2 Fe cations	Catalytic	1.14.16.5	Endoplasmic reticulum	Yes		ether lipid metabolic process [GO:0046485]; membrane lipid metabolic process [GO:0006643]; triglyceride biosynthetic process [GO:0019432]

14	Q12797	ASPH_HUMAN	ASPH BAH	Aspartyl/asparaginyl beta-hydroxylase (EC 1.14.11.16) (Aspartate beta-hydroxylase) (ASP beta-hydroxylase) (Peptide-aspartate beta-dioxygenase)	H679-H725	1 Fe cation	Catalytic	1.14.11.16	Endoplasmic reticulum	Yes	DISEASE: Facial dysmorphism, lens dislocation, anterior segment abnormalities, and spontaneous filtering blebs (FDLAB) [MIM:601552]: A syndrome characterized by dislocated crystalline lenses and anterior segment abnormalities in association with a distinctive facies involving flat cheeks and a beaked nose. Some affected individuals develop highly unusual non-traumatic conjunctival cysts (filtering blebs). {ECO:0000269|PubMed:24768550}. Note=The disease is caused by mutations affecting the gene represented in this entry.	activation of cysteine-type endopeptidase activity [GO:0097202]; activation of store-operated calcium channel activity [GO:0032237]; calcium ion transmembrane transport [GO:0070588]; cellular response to calcium ion [GO:0071277]; detection of calcium ion [GO:0005513]; face morphogenesis [GO:0060325]; ion transmembrane transport [GO:0034220]; limb morphogenesis [GO:0035108]; muscle contraction [GO:0006936]; negative regulation of cell proliferation [GO:0008285]; palate development [GO:0060021]; pattern specification process [GO:0007389]; peptidyl-aspartic acid hydroxylation [GO:0042264]; positive regulation of calcium ion transport into cytosol [GO:0010524]; positive regulation of intracellular protein transport [GO:0090316]; positive regulation of proteolysis [GO:0045862]; positive regulation of ryanodine-sensitive calcium-release channel activity [GO:0060316]; positive regulation of transcription, DNA-templated [GO:0045893]; regulation of cardiac conduction [GO:1903779]; regulation of cardiac muscle contraction by regulation of the release of sequestered calcium ion [GO:0010881]; regulation of cell communication by electrical coupling [GO:0010649]; regulation of inositol 1,4,5-trisphosphate-sensitive calcium-release channel activity [GO:0031585]; regulation of protein depolymerization [GO:1901879]; regulation of protein stability [GO:0031647]; regulation of release of sequestered calcium ion into cytosol by sarcoplasmic reticulum [GO:0010880]; regulation of ryanodine-sensitive calcium-release channel activity [GO:0060314]; response to ATP [GO:0033198]
15	Q6ICH7	ASPH2_HUMAN	ASPHD2	Aspartate beta-hydroxylase domain-containing protein 2 (EC 1.14.11.-)	H283-H328	1 Fe cation	Catalytic	1.14.11.-	Unknown	Yes		peptidyl-amino acid modification [GO:0018193]
16	Q9HAY6	BCDO1_HUMAN	BCO1 BCDO BCDO1 BCMO1	Beta,beta-carotene 15,15'-dioxygenase (EC 1.13.11.63) (Beta-carotene dioxygenase 1) (Beta-carotene oxygenase 1)	H172-H237-H308-H514	1 Fe cation	Catalytic	1.13.11.63	Unknown	No	DISEASE: Hypercarotenemia and vitamin A deficiency, autosomal dominant (ADHVAD) [MIM:115300]: A disorder characterized by increased serum beta-carotene, decreased conversion of beta-carotene to vitamin A and decreased serum vitamin A. {ECO:0000269|PubMed:17951468}. Note=The disease is caused by mutations affecting the gene represented in this entry.	beta-carotene metabolic process [GO:1901810]; retinal metabolic process [GO:0042574]; retinoid metabolic process [GO:0001523]; retinol metabolic process [GO:0042572]; vitamin A biosynthetic process [GO:0035238]
17	Q9BYV7	BCDO2_HUMAN	BCO2 BCDO2	Beta,beta-carotene 9',10'-oxygenase (EC 1.13.11.71) (B-diox-II) (Beta-carotene dioxygenase 2)	H226-H286-H357-H573	1 Fe cation	Catalytic	1.13.11.71	Mitochondrion	No		carotene catabolic process [GO:0016121]; carotene metabolic process [GO:0016119]; carotenoid metabolic process [GO:0016116]; oxidation-reduction process [GO:0055114]; regulation of mitochondrial membrane potential [GO:0051881]; regulation of reactive oxygen species metabolic process [GO:2000377]; retinal metabolic process [GO:0042574]; retinoic acid metabolic process [GO:0042573]; retinoid metabolic process [GO:0001523]; xanthophyll metabolic process [GO:0016122]
18	O75936	BODG_HUMAN	BBOX1 BBH BBOX	Gamma-butyrobetaine dioxygenase (EC 1.14.11.1) (Gamma-butyrobetaine hydroxylase) (Gamma-BBH) (Gamma-butyrobetaine,2-oxoglutarate dioxygenase)	H202-D204-H347	1 Fe cation	Catalytic	1.14.11.1	Cytoplasm	No		carnitine biosynthetic process [GO:0045329]
19	O95992	CH25H_HUMAN	CH25H	Cholesterol 25-hydroxylase (EC 1.14.99.38) (Cholesterol 25-monooxygenase) (h25OH)	H143-H147-H157-H161; H205-H238-H242-H243	2 Fe cations	Catalytic	1.14.99.38	Endoplasmic reticulum	Yes		B cell chemotaxis [GO:0035754]; bile acid biosynthetic process [GO:0006699]; cholesterol metabolic process [GO:0008203]; lipid metabolic process [GO:0006629]; sterol biosynthetic process [GO:0016126]
20	Q99807	COQ7_HUMAN	COQ7	5-demethoxyubiquinone hydroxylase, mitochondrial (DMQ hydroxylase) (EC 1.14.13.-) (Timing protein clk-1 homolog) (Ubiquinone biosynthesis monooxygenase COQ7)	E60-E90-H93-E178; E90-E142-E178-H181	2 Fe cations	Catalytic	1.14.13.-	Mitochondrion	Yes	DISEASE: Coenzyme Q10 deficiency, primary, 8 (COQ10D8) [MIM:616733]: An autosomal recessive disorder resulting from mitochondrial dysfunction and characterized by decreased levels of coenzyme Q10. Patients manifest neonatal lung hypoplasia, contractures, early infantile hypertension and cardiac hypertrophy, secondary to prenatal kidney dysplasia, with neonatal and infantile renal dysfunction. Clinical features also include progressive peripheral neuropathy, muscular hypotonia and atrophy, and mild psychomotor delay with hearing and visual impairment. {ECO:0000269|PubMed:26084283}. Note=The disease is caused by mutations affecting the gene represented in this entry.	negative regulation of transcription from RNA polymerase II promoter [GO:0000122]; positive regulation of transcription from RNA polymerase II promoter [GO:0045944]; regulation of reactive oxygen species metabolic process [GO:2000377]; ubiquinone biosynthetic process [GO:0006744]
21	D3DRM8	D3DRM8_HUMAN	hCG_2040046	Galactose-1-phosphate uridylyltransferase	E154-H253-H271-H273	1 Fe cation	Catalytic	2.7.7.12	Unknown	No		
22	Q9BU89	DOHH_HUMAN	DOHH HLRC1	Deoxyhypusine hydroxylase (hDOHH) (EC 1.14.99.29) (Deoxyhypusine dioxygenase) (Deoxyhypusine monooxygenase) (HEAT-like repeat-containing protein 1)	H56-E57-H89-E90; H207-E208-H240-E241	2 Fe cations	Catalytic	1.14.99.29	Unknown	No		peptidyl-lysine modification to peptidyl-hypusine [GO:0008612]

23	Q9GZT9	EGLN1_HUMAN	EGLN1 C1orf12 PNAS-118 PNAS-137	Egl nine homolog 1 (EC 1.14.11.29) (Hypoxia-inducible factor prolyl hydroxylase 2) (HIF-PH2) (HIF-prolyl hydroxylase 2) (HPH-2) (Prolyl hydroxylase domain-containing protein 2) (PHD2) (SM-20)	H313-D315-H374	1 Fe cation	Catalytic	1.14.11.29	Cytoplasm, Nucleus	No	DISEASE: Erythrocytosis, familial, 3 (ECYT3) [MIM:609820]: An autosomal dominant disorder characterized by increased serum red blood cell mass, elevated serum hemoglobin and hematocrit, and normal serum erythropoietin levels. {ECO:0000269|PubMed:16407130, ECO:0000269|PubMed:17579185}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cardiac muscle tissue morphogenesis [GO:0055008]; cellular iron ion homeostasis [GO:0006879]; heart trabecula formation [GO:0060347]; labyrinthine layer development [GO:0060711]; negative regulation of cAMP catabolic process [GO:0030821]; negative regulation of cyclic-nucleotide phosphodiesterase activity [GO:0051344]; negative regulation of sequence-specific DNA binding transcription factor activity [GO:0043433]; oxygen homeostasis [GO:0032364]; peptidyl-proline hydroxylation to 4-hydroxy-L-proline [GO:0018401]; positive regulation of transcription from RNA polymerase II promoter [GO:0045944]; regulation of angiogenesis [GO:0045765]; regulation of neuron death [GO:1901214]; regulation of transcription from RNA polymerase II promoter in response to hypoxia [GO:0061418]; response to hypoxia [GO:0001666]; response to nitric oxide [GO:0071731]; ventricular septum morphogenesis [GO:0060412]
24	Q96KS0	EGLN2_HUMAN	EGLN2 EIT6	Egl nine homolog 2 (EC 1.14.11.29) (Estrogen-induced tag 6) (HPH-3) (Hypoxia-inducible factor prolyl hydroxylase 1) (HIF-PH1) (HIF-prolyl hydroxylase 1) (HPH-1) (Prolyl hydroxylase domain-containing protein 1) (PHD1)	H297-D299-H358	1 Fe cation	Catalytic	1.14.11.29	Nucleus	No		cell redox homeostasis [GO:0045454]; intracellular estrogen receptor signaling pathway [GO:0030520]; peptidyl-proline hydroxylation to 4-hydroxy-L-proline [GO:0018401]; positive regulation of protein catabolic process [GO:0045732]; regulation of cell growth [GO:0001558]; regulation of neuron apoptotic process [GO:0043523]; regulation of transcription from RNA polymerase II promoter in response to hypoxia [GO:0061418]; response to hypoxia [GO:0001666]
25	Q9H6Z9	EGLN3_HUMAN	EGLN3	Egl nine homolog 3 (EC 1.14.11.29) (HPH-1) (Hypoxia-inducible factor prolyl hydroxylase 3) (HIF-PH3) (HIF-prolyl hydroxylase 3) (HPH-3) (Prolyl hydroxylase domain-containing protein 3) (PHD3)	H135-D137-H196	1 Fe cation	Catalytic	1.14.11.29	Cytoplasm, Nucleus	No		activation of cysteine-type endopeptidase activity involved in apoptotic process [GO:0006919]; apoptotic process [GO:0006915]; cellular response to DNA damage stimulus [GO:0006974]; peptidyl-proline hydroxylation to 4-hydroxy-L-proline [GO:0018401]; protein hydroxylation [GO:0018126]; regulation of cell proliferation [GO:0042127]; regulation of neuron apoptotic process [GO:0043523]; regulation of transcription from RNA polymerase II promoter in response to hypoxia [GO:0061418]; response to hypoxia [GO:0001666]
26	O95571	ETHE1_HUMAN	ETHE1 HSCO	Persulfide dioxygenase ETHE1, mitochondrial (EC 1.13.11.18) (Ethylmalonic encephalopathy protein 1) (Hepatoma subtracted clone one protein) (Sulfur dioxygenase ETHE1)	H79-H135-D154	1 Fe cation	Catalytic	1.13.11.18	Cytoplasm, Mitochondrion, Nucleus	No	DISEASE: Ethylmalonic encephalopathy (EE) [MIM:602473]: Autosomal recessive disorder characterized by neurodevelopmental delay and regression, recurrent petechiae, acrocyanosis, diarrhea, leading to death in the first decade of life. It is also associated with persistent lactic acidemia and ethylmalonic and methylsuccinic aciduria. {ECO:0000269|PubMed:14732903, ECO:0000269|PubMed:18593870, ECO:0000269|PubMed:23144459}. Note=The disease is caused by mutations affecting the gene represented in this entry.	glutathione metabolic process [GO:0006749]; hydrogen sulfide metabolic process [GO:0070813]; sulfide oxidation, using sulfide:quinone oxidoreductase [GO:0070221]
27	Q9UKA1	FBXL5_HUMAN	FBXL5 FBL4 FBL5 FLR1	F-box/LRR-repeat protein 5 (F-box and leucine-rich repeat protein 5) (F-box protein FBL4/FBL5) (p45SKP2-like protein)	H15-H57-E58-E61-E130; E61-H80-H126-E130	2 Fe cations	Substrate - regulation		Cytoplasm	No		iron ion homeostasis [GO:0055072]; positive regulation of cellular protein catabolic process [GO:1903364]; protein polyubiquitination [GO:0000209]; protein ubiquitination [GO:0016567]; SCF-dependent proteasomal ubiquitin-dependent protein catabolic process [GO:0031146]
28	Q9BXU8	FHL17_HUMAN	FTHL17	Ferritin heavy polypeptide-like 17 (Cancer/testis antigen 38) (CT38)	E28-D45-E50-E65-H66-E108-E135-Q142	Several Fe cations	Substrate - storage/transport		Unknown	No		intracellular sequestering of iron ion [GO:0006880]; iron ion transport [GO:0006826]
29	P0C7X4	FHL19_HUMAN	FTH1P19 FTHL19	Putative ferritin heavy polypeptide-like 19 (Ferritin heavy polypeptide 1 pseudogene 19)	D6-E13-E25-E28-E32-D95-D100	Several Fe cations	Substrate - storage/transport		Unknown	No		intracellular sequestering of iron ion [GO:0006880]; iron ion transport [GO:0006826]

30	Q16595	FRDA_HUMAN	FXN FRDA X25	Frataxin, mitochondrial (EC 1.16.3.1) (Friedreich ataxia protein) (Fxn) [Cleaved into: Frataxin intermediate form (i-FXN); Frataxin(56-210) (m56-FXN); Frataxin(78-210) (d-FXN) (m78-FXN); Frataxin mature form (Frataxin(81-210)) (m81-FXN)]	Unknown	1 Fe cation	Substrate - storage/transport	1.16.3.1	Cytoplasm, Mitochondrion	No	DISEASE: Friedreich ataxia (FRDA) [MIM:229300]: Autosomal recessive, progressive degenerative disease characterized by neurodegeneration and cardiomyopathy it is the most common inherited ataxia. The disorder is usually manifest before adolescence and is generally characterized by incoordination of limb movements, dysarthria, nystagmus, diminished or absent tendon reflexes, Babinski sign, impairment of position and vibratory senses, scoliosis, pes cavus, and hammer toe. In most patients, FRDA is due to GAA triplet repeat expansions in the first intron of the frataxin gene. But in some cases the disease is due to mutations in the coding region. {ECO:0000269|PubMed:10732799, ECO:0000269|PubMed:10874325, ECO:0000269|PubMed:19629184, ECO:0000269|PubMed:9150176, ECO:0000269|PubMed:9779809, ECO:0000269|PubMed:9989622, ECO:0000269|Ref.35, ECO:0000269|Ref.7, ECO:0000269|Ref.8}. Note=The disease is caused by mutations affecting the gene represented in this entry.	adult walking behavior [GO:0007628]; aerobic respiration [GO:0009060]; cellular iron ion homeostasis [GO:0006879]; cellular response to hydrogen peroxide [GO:0070301]; embryo development ending in birth or egg hatching [GO:0009792]; heme biosynthetic process [GO:0006783]; ion transport [GO:0006811]; iron incorporation into metallo-sulfur cluster [GO:0018283]; mitochondrion organization [GO:0007005]; negative regulation of apoptotic process [GO:0043066]; negative regulation of multicellular organism growth [GO:0040015]; negative regulation of organ growth [GO:0046621]; negative regulation of release of cytochrome c from mitochondria [GO:0090201]; oxidative phosphorylation [GO:0006119]; positive regulation of aconitate hydratase activity [GO:1904234]; positive regulation of catalytic activity [GO:0043085]; positive regulation of cell growth [GO:0030307]; positive regulation of cell proliferation [GO:0008284]; positive regulation of lyase activity [GO:0051349]; positive regulation of succinate dehydrogenase activity [GO:1904231]; proprioception [GO:0019230]; protein autoprocessing [GO:0016540]; regulation of ferrochelatase activity [GO:0010722]; response to iron ion [GO:0010039]; small molecule metabolic process [GO:0044281]
31	P02794	FRIH_HUMAN	FTH1 FTH FTHL6 OK/SW-cl.84 PIG15	Ferritin heavy chain (Ferritin H subunit) (EC 1.16.3.1) (Cell proliferation-inducing gene 15 protein) [Cleaved into: Ferritin heavy chain, N-terminally processed]	E28-D43-H58-Q59-E62-E63-E65-H66-E108-D132-Q142	Several Fe cations	Substrate - storage/transport	1.16.3.1	Unknown	No	DISEASE: Hemochromatosis 5 (HFE5) [MIM:615517]: A disorder of iron metabolism characterized by iron overload. Excess iron is deposited in a variety of organs leading to their failure, and resulting in serious illnesses including cirrhosis, hepatomas, diabetes, cardiomyopathy, arthritis, and hypogonadotropic hypogonadism. Severe effects of the disease usually do not appear until after decades of progressive iron loading. {ECO:0000269|PubMed:11389486}. Note=The disease is caused by mutations affecting the gene represented in this entry. In a Japanese family affected by HFE5, a single point mutation has been detected in the iron-responsive element (IRE) in the 5'-UTR of FTH1 mRNA. This mutation leads to an increased binding affinity for iron regulatory protein and thereby to the efficient suppression of mRNA translation.	cellular iron ion homeostasis [GO:0006879]; immune response [GO:0006955]; intracellular sequestering of iron ion [GO:0006880]; iron ion import [GO:0097286]; negative regulation of cell proliferation [GO:0008285]; negative regulation of fibroblast proliferation [GO:0048147]; neutrophil degranulation [GO:0043312]
32	P02792	FRIL_HUMAN	FTL	Ferritin light chain (Ferritin L subunit)	D39-D41-E46-E54-E57-E58-E61-E64-D128-E131	Several Fe cations	Substrate - storage/transport		Unknown	No	DISEASE: Hereditary hyperferritinemia-cataract syndrome (HHCS) [MIM:600886]: An autosomal dominant disease characterized by elevated level of ferritin in serum and tissues, and early-onset bilateral cataract. {ECO:0000269|PubMed:19176363}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Neurodegeneration with brain iron accumulation 3 (NBIA3) [MIM:606159]: A neurodegenerative disorder associated with iron accumulation in the brain, primarily in the basal ganglia. It is characterized by a variety of neurological signs including parkinsonism, ataxia, corticospinal signs, mild non-progressive cognitive deficit and episodic psychosis. It is linked with decreased serum ferritin levels. {ECO:0000269|PubMed:16116125}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: L-ferritin deficiency (LFTD) [MIM:615604]: A condition characterized by low levels of ferritin in serum and tissues in the absence of other hematological symptoms. Seizures and mild neuropsychologic impairment may manifest in individuals with complete ferritin deficiency. {ECO:0000269|PubMed:23940258}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cellular iron ion homeostasis [GO:0006879]; intracellular sequestering of iron ion [GO:0006880]; iron ion homeostasis [GO:0055072]; iron ion transport [GO:0006826]; neutrophil degranulation [GO:0043312]
33	Q8N4E7	FTMT_HUMAN	FTMT	Ferritin, mitochondrial (EC 1.16.3.1)	E87-D104-H117-Q118-E121-E122-E124-H125-E167-D191-Q201-H233	Several Fe cations	Substrate - storage/transport	1.16.3.1	Mitochondrion	No		cellular iron ion homeostasis [GO:0006879]; intracellular sequestering of iron ion [GO:0006880]; iron ion transport [GO:0006826]; positive regulation of aconitate hydratase activity [GO:1904234]; positive regulation of cell proliferation [GO:0008284]; positive regulation of lyase activity [GO:0051349]; positive regulation of succinate dehydrogenase activity [GO:1904231]

34	Q9C0B1	FTO_HUMAN	FTO KIAA1752	Alpha-ketoglutarate-dependent dioxygenase FTO (EC 1.14.11.-) (Fat mass and obesity-associated protein)	H231-D233-H307	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No	DISEASE: Growth retardation, developmental delay, and facial dysmorphism (GDFD) [MIM:612938]: A severe polymalformation syndrome characterized by postnatal growth retardation, microcephaly, severe psychomotor delay, functional brain deficits and characteristic facial dysmorphism. In some patients, structural brain malformations, cardiac defects, genital anomalies, and cleft palate are observed. Early death occurs by the age of 3 years. {ECO:0000269|PubMed:19559399, ECO:0000269|PubMed:22002720, ECO:0000269|PubMed:26378117, ECO:0000269|PubMed:26697951}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Obesity (OBESITY) [MIM:601665]: A condition characterized by an increase of body weight beyond the limitation of skeletal and physical requirements, as the result of excessive accumulation of body fat. {ECO:0000269|PubMed:26287746}. Note=Disease susceptibility is associated with variations affecting the gene represented in this entry. A pathogenic intronic FTO variation (rs1421085) disrupts an evolutionarily conserved motif for ARID5B binding. Loss of ARID5B binding results in overexpression of two genes distal to FTO, IRX3 and IRX5. IRX3 and IRX5 overexpression shifts pre-adipocytes differentiation from brown to white fat cells, resulting in increased lipid storage and loss of mitochondrial thermogenesis. {ECO:0000269|PubMed:26287746}.	adipose tissue development [GO:0060612]; DNA dealkylation involved in DNA repair [GO:0006307]; DNA demethylation [GO:0080111]; oxidative demethylation [GO:0070989]; oxidative single-stranded DNA demethylation [GO:0035552]; oxidative single-stranded RNA demethylation [GO:0035553]; regulation of brown fat cell differentiation [GO:0090335]; regulation of lipid storage [GO:0010883]; regulation of multicellular organism growth [GO:0040014]; regulation of respiratory system process [GO:0044065]; regulation of white fat cell proliferation [GO:0070350]; RNA repair [GO:0042245]; temperature homeostasis [GO:0001659]
35	P07902	GALT_HUMAN	GALT	Galactose-1-phosphate uridylyltransferase (Gal-1-P uridylyltransferase) (EC 2.7.7.12) (UDP-glucose--hexose-1-phosphate uridylyltransferase)	H301-H319-H321	1 Fe cation	Catalytic	2.7.7.12	Unknown	No	DISEASE: Galactosemia (GALCT) [MIM:230400]: Inherited disorder of galactose metabolism that causes jaundice, cataracts, and mental retardation. {ECO:0000269|PubMed:10220154, ECO:0000269|PubMed:11754113, ECO:0000269|PubMed:11919338, ECO:0000269|PubMed:1373122, ECO:0000269|PubMed:1427861, ECO:0000269|PubMed:15841485, ECO:0000269|PubMed:1610789, ECO:0000269|PubMed:17041746, ECO:0000269|PubMed:17876724, ECO:0000269|PubMed:18956253, ECO:0000269|PubMed:1897530, ECO:0000269|PubMed:2011574, ECO:0000269|PubMed:22461411, ECO:0000269|PubMed:23022339, ECO:0000269|PubMed:25592817, ECO:0000269|PubMed:25614870, ECO:0000269|PubMed:7550229, ECO:0000269|PubMed:7887416, ECO:0000269|PubMed:7887417, ECO:0000269|PubMed:8112740, ECO:0000269|PubMed:8499924, ECO:0000269|PubMed:8598637, ECO:0000269|PubMed:8741038, ECO:0000269|PubMed:8869397, ECO:0000269|PubMed:8956044, ECO:0000269|PubMed:9222760}. Note=The disease is caused by mutations affecting the gene represented in this entry.	galactose catabolic process [GO:0019388]; galactose metabolic process [GO:0006012]; UDP-glucose catabolic process [GO:0006258]

36	P09211	GSTP1_HUMAN	GSTP1 FAEES3 GST3	Glutathione S-transferase P (EC 2.5.1.18) (GST class-pi) (GSTP1-1)	Y8	1 Fe cation	Regulation - catalysis	2.5.1.18	Cytoplasm, Mitochondrion, Nucleus	No		animal organ regeneration [GO:0031100]; cellular response to cell-matrix adhesion [GO:0071460]; cellular response to epidermal growth factor stimulus [GO:0071364]; cellular response to glucocorticoid stimulus [GO:0071385]; cellular response to insulin stimulus [GO:0032869]; cellular response to lipopolysaccharide [GO:0071222]; central nervous system development [GO:0007417]; common myeloid progenitor cell proliferation [GO:0035726]; glutathione derivative biosynthetic process [GO:1901687]; glutathione metabolic process [GO:0006749]; linoleic acid metabolic process [GO:0043651]; negative regulation of acute inflammatory response [GO:0002674]; negative regulation of apoptotic process [GO:0043066]; negative regulation of biosynthetic process [GO:0009890]; negative regulation of ERK1 and ERK2 cascade [GO:0070373]; negative regulation of extrinsic apoptotic signaling pathway [GO:2001237]; negative regulation of fibroblast proliferation [GO:0048147]; negative regulation of I-kappaB kinase/NF-kappaB signaling [GO:0043124]; negative regulation of interleukin-1 beta production [GO:0032691]; negative regulation of JUN kinase activity [GO:0043508]; negative regulation of leukocyte proliferation [GO:0070664]; negative regulation of MAPK cascade [GO:0043409]; negative regulation of MAP kinase activity [GO:0043407]; negative regulation of monocyte chemotactic protein-1 production [GO:0071638]; negative regulation of nitric-oxide synthase biosynthetic process [GO:0051771]; negative regulation of protein kinase activity [GO:0006469]; negative regulation of smooth muscle cell chemotaxis [GO:0071672]; negative regulation of stress-activated MAPK cascade [GO:0032873]; negative regulation of tumor necrosis factor-mediated signaling pathway [GO:0010804]; negative regulation of tumor necrosis factor production [GO:0032720]; negative regulation of vascular smooth muscle cell proliferation [GO:1904706]; neutrophil degranulation [GO:0043312]; nitric oxide storage [GO:0035732]; oligodendrocyte development [GO:0014003]; positive regulation of superoxide anion generation [GO:0032930]; regulation of ERK1 and ERK2 cascade [GO:0070372]; regulation of stress-activated MAPK cascade [GO:0032872]; response to amino acid [GO:0043200]; response to estradiol [GO:0032355]; response to ethanol [GO:0045471]; response to L-ascorbic acid [GO:0033591]; response to reactive oxygen species [GO:0000302]; xenobiotic metabolic process [GO:0006805]
37	O43593	HAIR_HUMAN	HR	Lysine-specific demethylase hairless (EC 1.14.11.-)	C1007-E1009-H1125	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No	DISEASE: Alopecia universalis congenita (ALUNC) [MIM:203655]: A rare disorder characterized by loss of hair from the entire body. No hair are present in hair follicles on skin biopsy. {ECO:0000269|PubMed:12406339, ECO:0000269|PubMed:24334705, ECO:0000269|PubMed:9445480, ECO:0000269|PubMed:9736769}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Atrichia with papular lesions (APL) [MIM:209500]: An autosomal recessive disease characterized by papillary lesions over most of the body and almost complete absence of hair. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Hypotrichosis 4 (HYPT4) [MIM:146550]: An autosomal dominant condition characterized by reduced amount of hair, alopecia, little or no eyebrows, eyelashes or body hair, and coarse, wiry, twisted hair in early childhood. {ECO:0000269|PubMed:19122663, ECO:0000269|PubMed:24961381}. Note=The disease is caused by mutations affecting the gene represented in this entry.	histone H3-K9 demethylation [GO:0033169]; negative regulation of transcription, DNA-templated [GO:0045892]; regulation of transcription, DNA-templated [GO:0006355]; transcription, DNA-templated [GO:0006351]
38	Q96EW2	HBAP1_HUMAN	HSPBAP1 PASS1	HSPB1-associated protein 1 (27 kDa heat shock protein-associated protein 1) (Protein associated with small stress protein 1)	H175-D177-H257	1 Fe cation	Catalytic		Cytoplasm	No	DISEASE: Note=A chromosomal aberration involving HSPBAP1 has been found in a family with renal carcinoma (PubMed:12939738). Translocation t(2;3)(q35;q21) with the putative pseudogene DIRC3 (PubMed:12939738). Produces a hybrid mRNA encoding a truncated HSPBAP1 lacking the first 36 amino acids (PubMed:12939738). {ECO:0000269|PubMed:12939738}.	

39	P22830	HEMH_HUMAN	FECH	Ferrochelatase, mitochondrial (EC 4.99.1.1) (Heme synthase) (Protoheme ferro-lyase)	H263-E343	1 Fe cation	Substrate - biosynthesis	4.99.1.1	Mitochondrion	Yes	DISEASE: Erythropoietic protoporphyria (EPP) [MIM:177000]: A form of porphyria. Porphyrias are inherited defects in the biosynthesis of heme, resulting in the accumulation and increased excretion of porphyrins or porphyrin precursors. They are classified as erythropoietic or hepatic, depending on whether the enzyme deficiency occurs in red blood cells or in the liver. Erythropoietic protoporphyria is marked by excessive protoporphyrin in erythrocytes, plasma, liver and feces, and by widely varying photosensitive skin changes ranging from a burning or pruritic sensation to erythema, edema and wheals. {ECO:0000269|PubMed:10942404, ECO:0000269|PubMed:11375302, ECO:0000269|PubMed:12063482, ECO:0000269|PubMed:12601550, ECO:0000269|PubMed:1376018, ECO:0000269|PubMed:15286165, ECO:0000269|PubMed:17196862, ECO:0000269|PubMed:1755842, ECO:0000269|PubMed:7910885, ECO:0000269|PubMed:8757534, ECO:0000269|PubMed:9211198, ECO:0000269|PubMed:9585598, ECO:0000269|PubMed:9740232}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cellular response to dexamethasone stimulus [GO:0071549]; generation of precursor metabolites and energy [GO:0006091]; heme biosynthetic process [GO:0006783]; protoporphyrinogen IX metabolic process [GO:0046501]; response to arsenic-containing substance [GO:0046685]; response to drug [GO:0042493]; response to ethanol [GO:0045471]; response to insecticide [GO:0017085]; response to lead ion [GO:0010288]; response to light stimulus [GO:0009416]; response to methylmercury [GO:0051597]; response to platinum ion [GO:0070541]
40	P81172	HEPC_HUMAN	HAMP HEPC LEAP1 UNQ487/PRO1003	Hepcidin (Liver-expressed antimicrobial peptide 1) (LEAP-1) (Putative liver tumor regressor) (PLTR) [Cleaved into: Hepcidin-25 (Hepc25); Hepcidin-20 (Hepc20)]	Unknown	Unknown	Substrate - regulation		Extracellular space	No	DISEASE: Hemochromatosis 2B (HFE2B) [MIM:613313]: A juvenile form of hemochromatosis, a disorder of iron metabolism with excess deposition of iron in a variety of organs leading to their failure, bronze skin pigmentation, hepatic cirrhosis, arthropathy and diabetes. The most common symptoms of juvenile hemochromatosis at presentation are hypogonadism and cardiomyopathy. {ECO:0000269|PubMed:12915468, ECO:0000269|PubMed:14630809, ECO:0000269|PubMed:14633868, ECO:0000269|PubMed:14670915, ECO:0000269|PubMed:15099344}. Note=The disease is caused by mutations affecting the gene represented in this entry.	acute-phase response [GO:0006953]; aging [GO:0007568]; antimicrobial humoral immune response mediated by antimicrobial peptide [GO:0061844]; cellular iron ion homeostasis [GO:0006879]; cellular response to bile acid [GO:1903413]; cellular response to interleukin-6 [GO:0071354]; cellular response to lipopolysaccharide [GO:0071222]; cellular response to tumor necrosis factor [GO:0071356]; cellular response to X-ray [GO:0071481]; defense response to bacterium [GO:0042742]; defense response to fungus [GO:0050832]; defense response to Gram-negative bacterium [GO:0050829]; defense response to Gram-positive bacterium [GO:0050830]; immune response [GO:0006955]; killing of cells of other organism [GO:0031640]; liver regeneration [GO:0097421]; multicellular organismal iron ion homeostasis [GO:0060586]; negative regulation of ferrous iron export [GO:1904039]; negative regulation of intestinal absorption [GO:1904479]; negative regulation of ion transmembrane transporter activity [GO:0032413]; negative regulation of iron channel activity [GO:1904255]; negative regulation of iron ion transmembrane transport [GO:0034760]; negative regulation of transcription by RNA polymerase II [GO:0000122]; positive regulation of cell growth involved in cardiac muscle cell development [GO:0061051]; positive regulation of protein polyubiquitination [GO:1902916]; positive regulation of receptor catabolic process [GO:2000646]; positive regulation of receptor internalization [GO:0002092]; response to erythropoietin [GO:0036017]; response to ethanol [GO:0045471]; response to iron ion [GO:0010039]; response to iron ion starvation [GO:1990641]; response to vitamin A [GO:0033189]; response to zinc ion [GO:0010043]
41	Q93099	HGD_HUMAN	HGD HGO	Homogentisate 1,2-dioxygenase (EC 1.13.11.5) (Homogentisate oxygenase) (Homogentisic acid oxidase) (Homogentisicase)	H335-E341-H371	1 Fe cation	Catalytic	1.13.11.5	Unknown	No	DISEASE: Alkaptonuria (AKU) [MIM:203500]: An autosomal recessive error of metabolism characterized by an increase in the level of homogentisic acid. The clinical manifestations are urine that turns dark on standing and alkalinization, black ochronotic pigmentation of cartilage and collagenous tissues, and spine arthritis. {ECO:0000269|PubMed:10205262, ECO:0000269|PubMed:10340975, ECO:0000269|PubMed:10482952, ECO:0000269|PubMed:10594001, ECO:0000269|PubMed:19862842, ECO:0000269|PubMed:21437689, ECO:0000269|PubMed:23353776, ECO:0000269|PubMed:23430897, ECO:0000269|PubMed:25681086, ECO:0000269|PubMed:8782815, ECO:0000269|PubMed:9154114, ECO:0000269|PubMed:9529363, ECO:0000269|PubMed:9630082}. Note=The disease is caused by mutations affecting the gene represented in this entry.	L-phenylalanine catabolic process [GO:0006559]; tyrosine catabolic process [GO:0006572]

42	Q9NWT6	HIF1N_HUMAN	HIF1AN FIH1	Hypoxia-inducible factor 1-alpha inhibitor (EC 1.14.11.30) (EC 1.14.11.n4) (Factor inhibiting HIF-1) (FIH-1) (Hypoxia-inducible factor asparagine hydroxylase)	H199-D201-H279	1 Fe cation	Catalytic	1.14.11.30; 1.14.11.n4	Cytoplasm, Nucleus	No		negative regulation of Notch signaling pathway [GO:0045746]; negative regulation of transcription from RNA polymerase II promoter in response to hypoxia [GO:0061428]; oxidation-reduction process [GO:0055114]; peptidyl-asparagine hydroxylation [GO:0042265]; peptidyl-aspartic acid hydroxylation [GO:0042264]; peptidyl-histidine hydroxylation [GO:0036138]; positive regulation of myoblast differentiation [GO:0045663]; positive regulation of vasculogenesis [GO:2001214]; regulation of transcription from RNA polymerase II promoter in response to hypoxia [GO:0061418]; transcription, DNA-templated [GO:0006351]
43	Q8IWW8	HOT_HUMAN	ADHFE1 HMFT2263	Hydroxyacid-oxoacid transhydrogenase, mitochondrial (HOT) (EC 1.1.99.24) (Alcohol dehydrogenase iron-containing protein 1) (ADHFe1) (Fe-containing alcohol dehydrogenase)	D242-H246-H330-H357	1 Fe cation	Catalytic	1.1.99.24	Mitochondrion	No		2-oxoglutarate metabolic process [GO:0006103]; molecular hydrogen transport [GO:0015993]
44	Q96IR7	HPDL_HUMAN	HPDL GLOXD1	4-hydroxyphenylpyruvate dioxygenase-like protein (EC 1.13.-.-) (Glyoxalase domain-containing protein 1)	H163-H258-E339	1 Fe cation	Catalytic	1.13.-.-	Unknown	No		aromatic amino acid family metabolic process [GO:0009072]
45	P32754	HPPD_HUMAN	HPD PPD	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27) (4-hydroxyphenylpyruvic acid oxidase) (4HPPD) (HPD) (HPPDase)	H183-H266-E349	1 Fe cation	Catalytic	1.13.11.27	Unknown	No	DISEASE: Tyrosinemia 3 (TYRSN3) [MIM:276710]: An inborn error of metabolism characterized by elevations of tyrosine in the blood and urine, seizures and mild mental retardation. {ECO:0000269|PubMed:10942115, ECO:0000269|PubMed:11073718}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Hawkinsinuria (HAWK) [MIM:140350]: An inborn error of tyrosine metabolism characterized by failure to thrive, persistent metabolic acidosis, fine and sparse hair, and excretion of the unusual cyclic amino acid metabolite, hawkinsin, in the urine. {ECO:0000269|PubMed:11073718}. Note=The disease is caused by mutations affecting the gene represented in this entry.	L-phenylalanine catabolic process [GO:0006559]; tyrosine catabolic process [GO:0006572]
46	Q96NU7	HUTI_HUMAN	AMDHD1 HMFT1272	Probable imidazolonepropionase (EC 3.5.2.7) (Amidohydrolase domain-containing protein 1)	H87-H89-H260-D334	1 Fe or Zn cation	Catalytic - no redox	3.5.2.7	Unknown	No		histidine catabolic process [GO:0006548]; histidine catabolic process to glutamate and formamide [GO:0019556]; histidine catabolic process to glutamate and formate [GO:0019557]
47	Q15652	JHD2C_HUMAN	JMJD1C JHDM2C KIAA1380 TRIP8	Probable JmjC domain-containing histone demethylation protein 2C (EC 1.14.11.-) (Jumonji domain-containing protein 1C) (Thyroid receptor-interacting protein 8) (TR-interacting protein 8) (TRIP-8)	H2336-E2338-H2466	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No		blood coagulation [GO:0007596]; histone H3-K9 demethylation [GO:0033169]; regulation of transcription, DNA-templated [GO:0006355]; transcription, DNA-templated [GO:0006351]
48	Q9H9V9	JMJD4_HUMAN	JMJD4	JmjC domain-containing protein 4 (Jumonji domain-containing protein 4)	H235-D237-H315	1 Fe cation	Catalytic		Unknown	No		
49	Q6NYC1	JMJD6_HUMAN	JMJD6 KIAA0585 PTDSR	Bifunctional arginine demethylase and lysyl-hydroxylase JMJD6 (EC 1.14.11.-) (Histone arginine demethylase JMJD6) (JmjC domain-containing protein 6) (Jumonji domain-containing protein 6) (Lysyl-hydroxylase JMJD6) (Peptide-lysine 5-dioxygenase JMJD6) (Phosphatidylserine receptor) (Protein PTDSR)	H187-D189-H273	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No		cell surface receptor signaling pathway [GO:0007166]; erythrocyte development [GO:0048821]; heart development [GO:0007507]; histone H3-R2 demethylation [GO:0070078]; histone H4-R3 demethylation [GO:0070079]; kidney development [GO:0001822]; lung development [GO:0030324]; macrophage activation [GO:0042116]; mRNA processing [GO:0006397]; peptidyl-lysine hydroxylation to 5-hydroxy-L-lysine [GO:0018395]; recognition of apoptotic cell [GO:0043654]; regulation of mRNA splicing, via spliceosome [GO:0048024]; regulation of transcription, DNA-templated [GO:0006355]; retina development in camera-type eye [GO:0060041]; RNA splicing [GO:0008380]; sprouting angiogenesis [GO:0002040]; T cell differentiation in thymus [GO:0033077]; transcription, DNA-templated [GO:0006351]
50	P0C870	JMJD7_HUMAN	JMJD7	JmjC domain-containing protein 7 (Jumonji domain-containing protein 7)	H178-D180-H277	1 Fe cation	Catalytic		Unknown	No		
51	Q96S16	JMJD8_HUMAN	JMJD8 C16orf20 PP14397	JmjC domain-containing protein 8 (Jumonji domain-containing protein 8)	H249-H251-H318	1 Fe cation	Catalytic		Unknown	No		
52	Q9P272	K1456_HUMAN	KIAA1456 C8orf79	Probable tRNA methyltransferase 9-like protein (TRM9L) (EC 2.1.1.-)	H112	1 Fe cation	Catalytic	2.1.1.-	Unknown	No		tRNA modification [GO:0006400]; tRNA wobble uridine modification [GO:0002098]
53	Q9Y2K7	KDM2A_HUMAN	KDM2A CXXC8 FBL7 FBXL11 JHDM1A KIAA1004	Lysine-specific demethylase 2A (EC 1.14.11.27) (CXXC-type zinc finger protein 8) (F-box and leucine-rich repeat protein 11) (F-box protein FBL7) (F-box protein Lilina) (F-box/LRR-repeat protein 11) (JmjC domain-containing histone demethylation protein 1A) ([Histone-H3]-lysine-36 demethylase 1A)	H212-D214-Y222-H284	1 Fe cation	Catalytic	1.14.11.27	Nucleus	No		double-strand break repair via nonhomologous end joining [GO:0006303]; histone H3-K36 demethylation [GO:0070544]; regulation of transcription, DNA-templated [GO:0006355]; transcription, DNA-templated [GO:0006351]

54	Q8NHM5	KDM2B_HUMAN	KDM2B CXXC2 FBL10 FBXL10 JHDM1B PCCX2	Lysine-specific demethylase 2B (EC 1.14.11.27) (CXXC-type zinc finger protein 2) (F-box and leucine-rich repeat protein 10) (F-box protein FBL10) (F-box/LRR-repeat protein 10) (JmjC domain-containing histone demethylation protein 1B) (Jumonji domain-containing EMSY-interactor methyltransferase motif protein) (Protein JEMMA) (Protein-containing CXXC domain 2) ([Histone-H3]-lysine-36 demethylase 1B)	H242-D244-H314	1 Fe cation	Catalytic	1.14.11.27	Nucleus	No		embryonic camera-type eye morphogenesis [GO:0048596]; forebrain development [GO:0030900]; fourth ventricle development [GO:0021592]; hindbrain development [GO:0030902]; histone H2A monoubiquitination [GO:0035518]; initiation of neural tube closure [GO:0021993]; lateral ventricle development [GO:0021670]; midbrain development [GO:0030901]; midbrain-hindbrain boundary morphogenesis [GO:0021555]; negative regulation of neural precursor cell proliferation [GO:2000178]; negative regulation of neuron apoptotic process [GO:0043524]; negative regulation of transcription from RNA polymerase II promoter [GO:0000122]; positive regulation of cell growth [GO:0030307]; positive regulation of stem cell population maintenance [GO:1902459]; spermatogenesis [GO:0007283]; third ventricle development [GO:0021678]; transcription, DNA-templated [GO:0006351]
55	Q9Y4C1	KDM3A_HUMAN	KDM3A JHDM2A JMJD1 JMJD1A KIAA0742 TSGA	Lysine-specific demethylase 3A (EC 1.14.11.-) (JmjC domain-containing histone demethylation protein 2A) (Jumonji domain-containing protein 1A)	H1120-D1122-H1249	1 Fe cation	Catalytic	1.14.11.-	Cytoplasm, Nucleus	No		androgen receptor signaling pathway [GO:0030521]; formaldehyde biosynthetic process [GO:0046293]; histone H3-K9 demethylation [GO:0033169]; histone H3-K9 dimethylation [GO:0036123]; hormone-mediated signaling pathway [GO:0009755]; negative regulation of histone H3-K9 methylation [GO:0051573]; positive regulation of transcription, DNA-templated [GO:0045893]; positive regulation of transcription from RNA polymerase II promoter [GO:0045944]; regulation of stem cell differentiation [GO:2000736]; regulation of stem cell population maintenance [GO:2000036]; spermatid nucleus elongation [GO:0007290]; transcription, DNA-templated [GO:0006351]
56	Q7LBC6	KDM3B_HUMAN	KDM3B C5orf7 JHDM2B JMJD1B KIAA1082	Lysine-specific demethylase 3B (EC 1.14.11.-) (JmjC domain-containing histone demethylation protein 2B) (Jumonji domain-containing protein 1B) (Nuclear protein 5qNCA)	H1604-H1689	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No		histone H3-K9 demethylation [GO:0033169]; regulation of transcription, DNA-templated [GO:0006355]; response to cisplatin [GO:0072718]; transcription, DNA-templated [GO:0006351]
57	O75164	KDM4A_HUMAN	KDM4A JHDM3A JMJD2 JMJD2A KIAA0677	Lysine-specific demethylase 4A (EC 1.14.11.-) (JmjC domain-containing histone demethylation protein 3A) (Jumonji domain-containing protein 2A)	H188-E190-H276	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No		cardiac muscle hypertrophy in response to stress [GO:0014898]; histone demethylation [GO:0016577]; negative regulation of astrocyte differentiation [GO:0048712]; negative regulation of autophagy [GO:0010507]; negative regulation of cell death [GO:0060548]; negative regulation of gene expression [GO:0010629]; negative regulation of histone H3-K9 trimethylation [GO:1900113]; negative regulation of transcription, DNA-templated [GO:0045892]; positive regulation of gene expression [GO:0010628]; positive regulation of neuron differentiation [GO:0045666]; response to nutrient levels [GO:0031667]; transcription, DNA-templated [GO:0006351]; viral process [GO:0016032]
58	O94953	KDM4B_HUMAN	KDM4B JHDM3B JMJD2B KIAA0876	Lysine-specific demethylase 4B (EC 1.14.11.-) (JmjC domain-containing histone demethylation protein 3B) (Jumonji domain-containing protein 2B)	H189-E191-H277	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No		regulation of transcription, DNA-templated [GO:0006355]; transcription, DNA-templated [GO:0006351]
59	Q9H3R0	KDM4C_HUMAN	KDM4C GASC1 JHDM3C JMJD2C KIAA0780	Lysine-specific demethylase 4C (EC 1.14.11.-) (Gene amplified in squamous cell carcinoma 1 protein) (GASC-1 protein) (JmjC domain-containing histone demethylation protein 3C) (Jumonji domain-containing protein 2C)	H190-E192-H278	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No		blastocyst formation [GO:0001825]; histone H3-K9 demethylation [GO:0033169]; negative regulation of histone H3-K9 trimethylation [GO:1900113]; positive regulation of cell proliferation [GO:0008284]; positive regulation of gene expression [GO:0010628]; positive regulation of neuron differentiation [GO:0045666]; regulation of stem cell differentiation [GO:2000736]; regulation of stem cell population maintenance [GO:2000036]; regulation of transcription from RNA polymerase II promoter [GO:0006357]; transcription, DNA-templated [GO:0006351]
60	Q6B0I6	KDM4D_HUMAN	KDM4D JHDM3D JMJD2D	Lysine-specific demethylase 4D (EC 1.14.11.-) (JmjC domain-containing histone demethylation protein 3D) (Jumonji domain-containing protein 2D)	H192-E194-H280	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No		cellular response to ionizing radiation [GO:0071479]; double-strand break repair via homologous recombination [GO:0000724]; histone H3-K9 demethylation [GO:0033169]; negative regulation of histone H3-K9 trimethylation [GO:1900113]; positive regulation of chromatin binding [GO:0035563]; positive regulation of double-strand break repair via nonhomologous end joining [GO:2001034]; regulation of protein phosphorylation [GO:0001932]; regulation of transcription, DNA-templated [GO:0006355]; transcription, DNA-templated [GO:0006351]
61	B2RXH2	KDM4E_HUMAN	KDM4E KDM4DL	Lysine-specific demethylase 4E (EC 1.14.11.-) (KDM4D-like protein) (Lysine-specific demethylase 4D-like)	H189-E191-H277	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No		covalent chromatin modification [GO:0016569]; regulation of transcription, DNA-templated [GO:0006355]; transcription, DNA-templated [GO:0006351]

62	P29375	KDM5A_HUMAN	KDM5A JARID1A RBBP2 RBP2	Lysine-specific demethylase 5A (EC 1.14.11.-) (Histone demethylase JARID1A) (Jumonji/ARID domain-containing protein 1A) (Retinoblastoma-binding protein 2) (RBBP-2)	H483-E485-H571	1 Fe cation	Catalytic	1.14.11.-	Mitochondrion, Nucleus	No		circadian regulation of gene expression [GO:0032922]; histone H3-K4 demethylation [GO:0034720]; male gonad development [GO:0008584]; negative regulation of histone deacetylase activity [GO:1901726]; negative regulation of transcription from RNA polymerase II promoter [GO:0000122]; positive regulation of transcription, DNA-templated [GO:0045893]; regulation of sequence-specific DNA binding transcription factor activity [GO:0051090]; spermatogenesis [GO:0007283]; transcription from RNA polymerase II promoter [GO:0006366]
63	Q9UGL1	KDM5B_HUMAN	KDM5B JARID1B PLU1 RBBP2H1	Lysine-specific demethylase 5B (EC 1.14.11.-) (Cancer/testis antigen 31) (CT31) (Histone demethylase JARID1B) (Jumonji/ARID domain-containing protein 1B) (PLU-1) (Retinoblastoma-binding protein 2 homolog 1) (RBP2-H1)	H499-E501-H587	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No		branching involved in mammary gland duct morphogenesis [GO:0060444]; cellular response to fibroblast growth factor stimulus [GO:0044344]; histone H3-K4 demethylation [GO:0034720]; lens fiber cell differentiation [GO:0070306]; mammary duct terminal end bud growth [GO:0060763]; negative regulation of transcription, DNA-templated [GO:0045892]; positive regulation of gene expression [GO:0010628]; positive regulation of mammary gland epithelial cell proliferation [GO:0033601]; post-embryonic development [GO:0009791]; regulation of estradiol secretion [GO:2000864]; regulation of transcription from RNA polymerase II promoter [GO:0006357]; response to fungicide [GO:0060992]; rhythmic process [GO:0048511]; single fertilization [GO:0007338]; transcription, DNA-templated [GO:0006351]; uterus morphogenesis [GO:0061038]
64	P41229	KDM5C_HUMAN	KDM5C DXS1272E JARID1C SMCX XE169	Lysine-specific demethylase 5C (EC 1.14.11.-) (Histone demethylase JARID1C) (Jumonji/ARID domain-containing protein 1C) (Protein SmcX) (Protein Xe169)	H514-E516-H602	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No	DISEASE: Mental retardation, X-linked, syndromic, Claes-Jensen type (MRXSCJ) [MIM:300534]: A disorder characterized by significantly below average general intellectual functioning associated with impairments in adaptive behavior and manifested during the developmental period. MRXSCJ patients manifest mental retardation associated with variable features such as slowly progressive spastic paraplegia, seizures, facial dysmorphism. {ECO:0000269|PubMed:15586325, ECO:0000269|PubMed:16538222, ECO:0000269|PubMed:16541399, ECO:0000269|PubMed:17320160, ECO:0000269|PubMed:17468742, ECO:0000269|PubMed:23356856, ECO:0000269|PubMed:25666439}. Note=The disease is caused by mutations affecting the gene represented in this entry.	histone H3-K4 demethylation [GO:0034720]; negative regulation of transcription, DNA-templated [GO:0045892]; response to toxic substance [GO:0009636]; rhythmic process [GO:0048511]; transcription, DNA-templated [GO:0006351]
65	Q9BY66	KDM5D_HUMAN	KDM5D HY HYA JARID1D KIAA0234 SMCY	Lysine-specific demethylase 5D (EC 1.14.11.-) (Histocompatibility Y antigen) (H-Y) (Histone demethylase JARID1D) (Jumonji/ARID domain-containing protein 1D) (Protein SmcY)	H504-E506-H592	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No		histone H3-K4 demethylation [GO:0034720]; regulation of androgen receptor signaling pathway [GO:0060765]; regulation of transcription, DNA-templated [GO:0006355]; T cell antigen processing and presentation [GO:0002457]; transcription, DNA-templated [GO:0006351]
66	O15550	KDM6A_HUMAN	KDM6A UTX	Lysine-specific demethylase 6A (EC 1.14.11.-) (Histone demethylase UTX) (Ubiquitously-transcribed TPR protein on the X chromosome) (Ubiquitously-transcribed X chromosome tetratricopeptide repeat protein)	H1146-E1148-H1226	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No	DISEASE: Kabuki syndrome 2 (KABUK2) [MIM:300867]: A congenital mental retardation syndrome with additional features, including postnatal dwarfism, a peculiar facies characterized by long palpebral fissures with eversion of the lateral third of the lower eyelids, a broad and depressed nasal tip, large prominent earlobes, a cleft or high-arched palate, scoliosis, short fifth finger, persistence of fingerpads, radiographic abnormalities of the vertebrae, hands, and hip joints, and recurrent otitis media in infancy. {ECO:0000269|PubMed:22197486}. Note=The disease is caused by mutations affecting the gene represented in this entry.	canonical Wnt signaling pathway [GO:0060070]; cardiovascular system development [GO:0072358]; heart morphogenesis [GO:0003007]; histone H3-K4 methylation [GO:0051568]; in utero embryonic development [GO:0001701]; mesodermal cell differentiation [GO:0048333]; multicellular organism growth [GO:0035264]; neural tube closure [GO:0001843]; notochord morphogenesis [GO:0048570]; positive regulation of gene expression [GO:0010628]; respiratory system process [GO:0003016]; somite rostral/caudal axis specification [GO:0032525]
67	O15054	KDM6B_HUMAN	KDM6B JMJD3 KIAA0346	Lysine-specific demethylase 6B (EC 1.14.11.-) (JmjC domain-containing protein 3) (Jumonji domain-containing protein 3) (Lysine demethylase 6B)	H1390-E1392-H1470	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No		cardiac muscle cell differentiation [GO:0055007]; cell fate commitment [GO:0045165]; cellular response to hydrogen peroxide [GO:0070301]; endothelial cell differentiation [GO:0045446]; hippocampus development [GO:0021766]; inflammatory response to antigenic stimulus [GO:0002437]; mesodermal cell differentiation [GO:0048333]; positive regulation of transcription from RNA polymerase II promoter [GO:0045944]; response to activity [GO:0014823]; response to fungicide [GO:0060992]
68	Q6ZMT4	KDM7A_HUMAN	KDM7A JHDM1D KDM7 KIAA1718	Lysine-specific demethylase 7A (EC 1.14.11.-) (JmjC domain-containing histone demethylation protein 1D) (Lysine-specific demethylase 7)	H282-D284-Y292-H354	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No		histone H3-K27 demethylation [GO:0071557]; histone H3-K36 demethylation [GO:0070544]; histone H3-K9 demethylation [GO:0033169]; histone H4-K20 demethylation [GO:0035574]; midbrain development [GO:0030901]; positive regulation of transcription, DNA-templated [GO:0045893]; transcription, DNA-templated [GO:0006351]

69	Q8N371	KDM8_HUMAN	KDM8 JMJD5	Lysine-specific demethylase 8 (EC 1.14.11.27) (JmjC domain-containing protein 5) (Jumonji domain-containing protein 5)	H321-D323-H400	1 Fe cation	Catalytic	1.14.11.27	Nucleus	No		G2/M transition of mitotic cell cycle [GO:0000086]; histone H3-K36 demethylation [GO:0070544]; positive regulation of transcription, DNA-templated [GO:0045893]; transcription, DNA-templated [GO:0006351]
70	P18054	LOX12_HUMAN	ALOX12 12LO LOG12	Arachidonate 12-lipoxygenase, 12S-type (12S-LOX) (12S-lipoxygenase) (EC 1.13.11.31) (Lipoxin synthase 12-LO) (EC 3.3.2.-) (Platelet-type lipoxygenase 12)	H360-H365-H540	1 Fe cation	Catalytic	1.13.11.31; 3.3.2.-	Cytoplasm	Yes	DISEASE: Esophageal cancer (ESCR) [MIM:133239]: A malignancy of the esophagus. The most common types are esophageal squamous cell carcinoma and adenocarcinoma. Cancer of the esophagus remains a devastating disease because it is usually not detected until it has progressed to an advanced incurable stage. {ECO:0000269|PubMed:17460548}. Note=Disease susceptibility may be associated with variations affecting the gene represented in this entry. Gln at position 261 may confer interindividual susceptibility to esophageal cancer (PubMed:17460548). {ECO:0000269|PubMed:17460548}.; DISEASE: Colorectal cancer (CRC) [MIM:114500]: A complex disease characterized by malignant lesions arising from the inner wall of the large intestine (the colon) and the rectum. Genetic alterations are often associated with progression from premalignant lesion (adenoma) to invasive adenocarcinoma. Risk factors for cancer of the colon and rectum include colon polyps, long-standing ulcerative colitis, and genetic family history. {ECO:0000269|PubMed:17151091}. Note=Disease susceptibility may be associated with variations affecting the gene represented in this entry. Gln at position 261 may confer interindividual susceptibility to colorectal cancer (PubMed:17460548). {ECO:0000269|PubMed:17460548}.	aging [GO:0007568]; arachidonic acid metabolic process [GO:0019369]; cellular response to lipid [GO:0071396]; establishment of skin barrier [GO:0061436]; fatty acid oxidation [GO:0019395]; hepoxilin biosynthetic process [GO:0051122]; hepoxilin metabolic process [GO:0051121]; leukotriene A4 metabolic process [GO:1901751]; linoleic acid metabolic process [GO:0043651]; lipoxin A4 biosynthetic process [GO:2001303]; lipoxin B4 biosynthetic process [GO:2001306]; lipoxin metabolic process [GO:2001300]; lipoxygenase pathway [GO:0019372]; movement of cell or subcellular component [GO:0006928]; negative regulation of apoptotic process [GO:0043066]; negative regulation of muscle cell apoptotic process [GO:0010656]; negative regulation of platelet aggregation [GO:0090331]; positive regulation of angiogenesis [GO:0045766]; positive regulation of blood vessel diameter [GO:0097755]; positive regulation of cell adhesion [GO:0045785]; positive regulation of cell growth [GO:0030307]; positive regulation of cell migration [GO:0030335]; positive regulation of cell proliferation [GO:0008284]; positive regulation of cysteine-type endopeptidase activity involved in apoptotic process [GO:0043280]; positive regulation of endothelial cell differentiation [GO:0045603]; positive regulation of endothelial cell migration [GO:0010595]; positive regulation of gene expression [GO:0010628]; positive regulation of mitochondrial depolarization [GO:0051901]; positive regulation of smooth muscle cell proliferation [GO:0048661]; reactive oxygen species metabolic process [GO:0072593]; superoxide anion generation [GO:0042554]
71	P16050	LOX15_HUMAN	ALOX15 LOG15	Arachidonate 15-lipoxygenase (15-LOX) (15-LOX-1) (EC 1.13.11.33) (12/15-lipoxygenase) (Arachidonate 12-lipoxygenase, leukocyte-type) (12-LOX) (EC 1.13.11.31) (Arachidonate omega-6 lipoxygenase)	H360-H365-H540	1 Fe cation	Catalytic	1.13.11.33; 1.13.11.31	Cytoplasm, Cell membrane	Yes	DISEASE: Note=Disease susceptibility may be associated with variations affecting the gene represented in this entry. Met at position 560 may confer interindividual susceptibility to coronary artery disease (CAD) (PubMed:17959182). {ECO:0000269|PubMed:17959182}.	apoptotic cell clearance [GO:0043277]; arachidonic acid metabolic process [GO:0019369]; bone mineralization [GO:0030282]; cellular response to calcium ion [GO:0071277]; cellular response to interleukin-13 [GO:0035963]; hepoxilin biosynthetic process [GO:0051122]; inflammatory response [GO:0006954]; leukotriene metabolic process [GO:0006691]; lipoxin A4 biosynthetic process [GO:2001303]; lipoxygenase pathway [GO:0019372]; negative regulation of adaptive immune response [GO:0002820]; ossification [GO:0001503]; phosphatidylethanolamine biosynthetic process [GO:0006646]; positive regulation of actin filament polymerization [GO:0030838]; positive regulation of cell-substrate adhesion [GO:0010811]; positive regulation of ERK1 and ERK2 cascade [GO:0070374]; positive regulation of heterotypic cell-cell adhesion [GO:0034116]; regulation of engulfment of apoptotic cell [GO:1901074]; regulation of peroxisome proliferator activated receptor signaling pathway [GO:0035358]; response to endoplasmic reticulum stress [GO:0034976]; wound healing [GO:0042060]
72	P09917	LOX5_HUMAN	ALOX5 LOG5	Arachidonate 5-lipoxygenase (5-LO) (5-lipoxygenase) (EC 1.13.11.34)	H368-H373-H551	1 Fe cation	Catalytic	1.13.11.34	Cytoplasm, Nucleus	Yes		leukotriene biosynthetic process [GO:0019370]; leukotriene metabolic process [GO:0006691]; leukotriene production involved in inflammatory response [GO:0002540]; lipoxin metabolic process [GO:2001300]; lipoxygenase pathway [GO:0019372]; neutrophil degranulation [GO:0043312]

73	Q9BYJ1	LOXE3_HUMAN	ALOXE3	Hydroperoxide isomerase ALOXE3 (EC 5.4.4.7) (Epidermis-type lipoxygenase 3) (Epidermal LOX-3) (e-LOX-3) (eLOX-3) (Hydroperoxy icosatetraenoate dehydratase) (EC 4.2.1.152)	H408-H413-H588	1 Fe cation	Catalytic	5.4.4.7; 4.2.1.152	Cytoplasm	No	DISEASE: Ichthyosis, congenital, autosomal recessive 3 (ARCI3) [MIM:606545]: A form of autosomal recessive congenital ichthyosis, a disorder of keratinization with abnormal differentiation and desquamation of the epidermis, resulting in abnormal skin scaling over the whole body. The main skin phenotypes are lamellar ichthyosis (LI) and non-bullous congenital ichthyosiform erythroderma (NCIE), although phenotypic overlap within the same patient or among patients from the same family can occur. Lamellar ichthyosis is a condition often associated with an embedment in a collodion-like membrane at birth; skin scales later develop, covering the entire body surface. Non-bullous congenital ichthyosiform erythroderma characterized by fine whitish scaling on an erythrodermal background; larger brownish scales are present on the buttocks, neck and legs. {ECO:0000269|PubMed:11773004, ECO:0000269|PubMed:15629692, ECO:0000269|PubMed:16116617, ECO:0000269|PubMed:19131948, ECO:0000269|PubMed:19890349}. Note=The disease is caused by mutations affecting the gene represented in this entry.	arachidonic acid metabolic process [GO:0019369]; ceramide biosynthetic process [GO:0046513]; establishment of skin barrier [GO:0061436]; fat cell differentiation [GO:0045444]; hepoxilin biosynthetic process [GO:0051122]; linoleic acid metabolic process [GO:0043651]; lipoxygenase pathway [GO:0019372]; peroxisome proliferator activated receptor signaling pathway [GO:0035357]; sensory perception of pain [GO:0019233]; sphingolipid metabolic process [GO:0006665]
74	O75342	LX12B_HUMAN	ALOX12B	Arachidonate 12-lipoxygenase, 12R-type (12R-LOX) (12R-lipoxygenase) (EC 1.13.11.-) (Epidermis-type lipoxygenase 12)	H398-H403-H578	1 Fe cation	Catalytic	1.13.11.-	Cytoplasm	No	DISEASE: Ichthyosis, congenital, autosomal recessive 2 (ARCI2) [MIM:242100]: A form of autosomal recessive congenital ichthyosis, a disorder of keratinization with abnormal differentiation and desquamation of the epidermis, resulting in abnormal skin scaling over the whole body. The main skin phenotypes are lamellar ichthyosis (LI) and non-bullous congenital ichthyosiform erythroderma (NCIE), although phenotypic overlap within the same patient or among patients from the same family can occur. Lamellar ichthyosis is a condition often associated with an embedment in a collodion-like membrane at birth; skin scales later develop, covering the entire body surface. Non-bullous congenital ichthyosiform erythroderma characterized by fine whitish scaling on an erythrodermal background; larger brownish scales are present on the buttocks, neck and legs. {ECO:0000269|PubMed:11773004, ECO:0000269|PubMed:15629692, ECO:0000269|PubMed:16116617, ECO:0000269|PubMed:19131948, ECO:0000269|PubMed:19890349}. Note=The disease is caused by mutations affecting the gene represented in this entry.	arachidonic acid metabolic process [GO:0019369]; ceramide biosynthetic process [GO:0046513]; establishment of skin barrier [GO:0061436]; hepoxilin biosynthetic process [GO:0051122]; linoleic acid metabolic process [GO:0043651]; lipoxygenase pathway [GO:0019372]; oxidation-reduction process [GO:0055114]; positive regulation of gene expression [GO:0010628]; positive regulation of MAPK cascade [GO:0043410]; positive regulation of mucus secretion [GO:0070257]; protein lipidation [GO:0006497]; sphingolipid metabolic process [GO:0006665]
75	O15296	LX15B_HUMAN	ALOX15B	Arachidonate 15-lipoxygenase B (15-LOX-B) (EC 1.13.11.33) (15-lipoxygenase 2) (15-LOX-2) (Arachidonate 15-lipoxygenase type II) (Linoleate 13-lipoxygenase 15-LOb) (EC 1.13.11.-)	H373-H378-H553	1 Fe cation	Catalytic	1.13.11.33; 1.13.11.-	Nucleus	No		apoptotic process [GO:0006915]; arachidonic acid metabolic process [GO:0019369]; hepoxilin biosynthetic process [GO:0051122]; linoleic acid metabolic process [GO:0043651]; lipid metabolic process [GO:0006629]; lipoxygenase pathway [GO:0019372]; negative regulation of cell cycle [GO:0045786]; negative regulation of cell migration [GO:0030336]; negative regulation of cell proliferation [GO:0008285]; negative regulation of growth [GO:0045926]; positive regulation of chemokine secretion [GO:0090197]; positive regulation of keratinocyte differentiation [GO:0045618]; positive regulation of macrophage derived foam cell differentiation [GO:0010744]; positive regulation of peroxisome proliferator activated receptor signaling pathway [GO:0035360]; prostate gland development [GO:0030850]; regulation of epithelial cell differentiation [GO:0030856]
76	P53582	MAP11_HUMAN	METAP1 KIAA0094	Methionine aminopeptidase 1 (MAP 1) (MetAP 1) (EC 3.4.11.18) (Peptidase M 1)	D220-D231-H294-E327-E358	1 Divalent cation	Catalytic	3.4.11.18	Cytoplasm	No		N-terminal protein amino acid modification [GO:0031365]; peptidyl-methionine modification [GO:0018206]; platelet aggregation [GO:0070527]; regulation of rhodopsin mediated signaling pathway [GO:0022400]; regulation of translation [GO:0006417]
77	Q6UB28	MAP12_HUMAN	METAP1D MAP1D	Methionine aminopeptidase 1D, mitochondrial (MAP 1D) (MetAP 1D) (EC 3.4.11.18) (Methionyl aminopeptidase type 1D, mitochondrial) (Peptidase M 1D)	D178-D189-H252-E284-E315	1 Divalent cation	Catalytic	3.4.11.18	Mitochondrion	No		N-terminal protein amino acid modification [GO:0031365]; peptidyl-methionine modification [GO:0018206]
78	P50579	MAP2_HUMAN	METAP2 MNPEP P67EIF2	Methionine aminopeptidase 2 (MAP 2) (MetAP 2) (EC 3.4.11.18) (Initiation factor 2-associated 67 kDa glycoprotein) (p67) (p67eIF2) (Peptidase M)	D251-D262-H331-E364-E459	1 Divalent cation	Catalytic	3.4.11.18	Cytoplasm	No		N-terminal protein amino acid modification [GO:0031365]; peptidyl-methionine modification [GO:0018206]; protein processing [GO:0016485]; regulation of rhodopsin mediated signaling pathway [GO:0022400]
79	Q9NYZ2	MFRN1_HUMAN	SLC25A37 MFRN MSCP HT015	Mitoferrin-1 (Mitochondrial iron transporter 1) (Mitochondrial solute carrier protein) (Solute carrier family 25 member 37)	Unknown	Unknown	Substrate - transport		Mitochondrion	Yes		iron ion homeostasis [GO:0055072]; mitochondrial iron ion transport [GO:0048250]

80	Q96A46	MFRN2_HUMAN	SLC25A28 MFRN2 NPD016	Mitoferrin-2 (Mitochondrial RNA-splicing protein 3/4 homolog) (MRS3/4) (hMRS3/4) (Mitochondrial iron transporter 2) (Solute carrier family 25 member 28)	Unknown	Unknown	Substrate - transport		Mitochondrion	Yes		iron ion homeostasis [GO:0055072]; mitochondrial iron ion transport [GO:0048250]
81	Q9UGB7	MIOX_HUMAN	MIOX ALDRL6 KSP32 RSOR	Inositol oxygenase (EC 1.13.99.1) (Aldehyde reductase-like 6) (Kidney-specific protein 32) (Myo-inositol oxygenase) (MI oxygenase) (Renal-specific oxidoreductase)	H98-H123-D124-D253; D124-H194-H220	2 Fe cations	Catalytic	1.13.99.1	Cytoplasm	No		inositol catabolic process [GO:0019310]
82	O15442	MPPD1_HUMAN	MPPED1 C22orf1 FAM1A	Metallophosphoesterase domain-containing protein 1 (EC 3.1.-.-) (Adult brain protein 239) (239AB)	D97-H99-D118-H286; H245-H284-N149	2 Divalent cations	Catalytic	3.1.-.-	Unknown	No		
83	P49959	MRE11_HUMAN	MRE11 HNGS1 MRE11A	Double-strand break repair protein MRE11 (Double-strand break repair protein MRE11A) (Meiotic recombination 11 homolog 1) (MRE11 homolog 1) (Meiotic recombination 11 homolog A) (MRE11 homolog A)	D20-H22-D60	1 Fe cation	Catalytic		Nucleus	No	DISEASE: Ataxia-telangiectasia-like disorder 1 (ATLD1) [MIM:604391]: A rare disorder characterized by progressive cerebellar ataxia, dysarthria, abnormal eye movements, and absence of telangiectasia. ATLD patients show normal levels of total IgG, IgA and IgM, although there may be reduced levels of specific functional antibodies. At the cellular level, ATLD exhibits hypersensitivity to ionizing radiation and radioresistant DNA synthesis. {ECO:0000269|PubMed:10612394}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Note=Defects in MRE11 can be a cause of nephronophthisis-related ciliopathies (NPHP-RC), a group of recessive diseases that affect kidney, retina and brain. A homozygous truncating mutation MRE11 has been found in patients with cerebellar vermis hypoplasia, ataxia and dysarthria. {ECO:0000269|PubMed:22863007}.	cell proliferation [GO:0008283]; cellular response to DNA damage stimulus [GO:0006974]; DNA double-strand break processing [GO:0000729]; DNA duplex unwinding [GO:0032508]; DNA recombination [GO:0006310]; DNA repair [GO:0006281]; DNA replication [GO:0006260]; DNA synthesis involved in DNA repair [GO:0000731]; double-strand break repair [GO:0006302]; double-strand break repair via homologous recombination [GO:0000724]; double-strand break repair via nonhomologous end joining [GO:0006303]; intra-S DNA damage checkpoint [GO:0031573]; mitotic G2 DNA damage checkpoint [GO:0007095]; negative regulation of apoptotic process [GO:0043066]; negative regulation of DNA endoreduplication [GO:0032876]; positive regulation of kinase activity [GO:0033674]; positive regulation of protein autophosphorylation [GO:0031954]; positive regulation of telomere maintenance [GO:0032206]; positive regulation of type I interferon production [GO:0032481]; reciprocal meiotic recombination [GO:0007131]; regulation of mitotic recombination [GO:0000019]; regulation of signal transduction by p53 class mediator [GO:1901796]; sister chromatid cohesion [GO:0007062]; strand displacement [GO:0000732]; synapsis [GO:0007129]; telomere maintenance via telomerase [GO:0007004]; telomeric 3' overhang formation [GO:0031860]; viral process [GO:0016032]
84	Q15800	MSMO1_HUMAN	MSMO1 DESP4 ERG25 SC4MOL	Methylsterol monooxygenase 1 (EC 1.14.13.72) (C-4 methylsterol oxidase)	Unknown	1 Fe cation	Catalytic	1.14.13.72	Endoplasmic reticulum	Yes	DISEASE: Microcephaly, congenital cataract, and psoriasiform dermatitis (MCCPD) [MIM:616834]: An autosomal recessive inborn error of cholesterol metabolism characterized by accumulation of a large amount of methylsterols, particularly dimethylsterols, in affected individuals. Patients manifest psoriasiform dermatitis, arthralgias, congenital cataracts, microcephaly, and developmental delay. {ECO:0000269|PubMed:21285510, ECO:0000269|PubMed:24144731}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cholesterol biosynthetic process [GO:0006695]; fatty acid metabolic process [GO:0006631]; steroid metabolic process [GO:0008202]; sterol biosynthetic process [GO:0016126]
85	Q9BV57	MTND_HUMAN	ADI1 MTCBP1 HMFT1638	1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase (EC 1.13.11.54) (Acireductone dioxygenase (Fe(2+)-requiring)) (ARD) (Fe-ARD) (Membrane-type 1 matrix metalloproteinase cytoplasmic tail-binding protein 1) (MTCBP-1) (Submergence-induced protein-like factor) (Sip-L)	H88-H90-E94-H133	1 Fe cation	Catalytic	1.13.11.54	Cytoplasm, Cell membrane, Nucleus	Yes		L-methionine salvage from methylthioadenosine [GO:0019509]
86	P80188	NGAL_HUMAN	LCN2 HNL NGAL	Neutrophil gelatinase-associated lipocalin (NGAL) (25 kDa alpha-2-microglobulin-related subunit of MMP-9) (Lipocalin-2) (Oncogene 24p3) (Siderocalin LCN2) (p25)	Y126-K145-K154	Binds ferric siderophore	Substrate - transport		Extracellular space	No		antimicrobial humoral response [GO:0019730]; cellular iron ion homeostasis [GO:0006879]; cellular response to hydrogen peroxide [GO:0070301]; cellular response to interleukin-1 [GO:0071347]; cellular response to lipopolysaccharide [GO:0071222]; cellular response to nutrient levels [GO:0031669]; cellular response to tumor necrosis factor [GO:0071356]; extrinsic apoptotic signaling pathway in absence of ligand [GO:0097192]; innate immune response [GO:0045087]; ion transport [GO:0006811]; neutrophil degranulation [GO:0043312]; positive regulation of cell projection organization [GO:0031346]; positive regulation of gene expression [GO:0010628]; protein homotrimerization [GO:0070207]; response to drug [GO:0042493]; response to herbicide [GO:0009635]; response to mycotoxin [GO:0010046]; response to virus [GO:0009615]; siderophore transport [GO:0015891]
87	Q9GZT8	NIF3L_HUMAN	NIF3L1 ALS2CR1 MDS015 My018	NIF3-like protein 1 (Amyotrophic lateral sclerosis 2 chromosomal region candidate gene 1 protein)	H93-H339-E343	1 Fe cation	Catalytic		Cytoplasm, Nucleus	No		negative regulation of nucleic acid-templated transcription [GO:1903507]; neuron differentiation [GO:0030182]; positive regulation of transcription, DNA-templated [GO:0045893]

88	P49279	NRAM1_HUMAN	SLC11A1 LSH NRAMP NRAMP1	Natural resistance-associated macrophage protein 1 (NRAMP 1) (Solute carrier family 11 member 1)	Unknown	Unknown	Substrate - transport		Unknown	Yes		activation of protein kinase activity [GO:0032147]; antigen processing and presentation of peptide antigen [GO:0048002]; antimicrobial humoral response [GO:0019730]; cadmium ion transmembrane transport [GO:0070574]; cell redox homeostasis [GO:0045454]; cellular cadmium ion homeostasis [GO:0006876]; cellular iron ion homeostasis [GO:0006879]; defense response to bacterium [GO:0042742]; defense response to Gram-negative bacterium [GO:0050829]; defense response to protozoan [GO:0042832]; divalent metal ion export [GO:0070839]; inflammatory response [GO:0006954]; interleukin-2 production [GO:0032623]; interleukin-3 production [GO:0032632]; iron ion homeostasis [GO:0055072]; iron ion transport [GO:0006826]; L-arginine import [GO:0043091]; macrophage activation [GO:0042116]; manganese ion transport [GO:0006828]; MHC class II biosynthetic process [GO:0045342]; mRNA stabilization [GO:0048255]; multicellular organismal iron ion homeostasis [GO:0060586]; negative regulation of cytokine production [GO:0001818]; neutrophil degranulation [GO:0043312]; nitrite transport [GO:0015707]; phagocytosis [GO:0006909]; positive regulation of cytokine production [GO:0001819]; positive regulation of dendritic cell antigen processing and presentation [GO:0002606]; positive regulation of gene expression [GO:0010628]; positive regulation of interferon-gamma production [GO:0032729]; positive regulation of phagocytosis [GO:0050766]; positive regulation of T-helper 1 type immune response [GO:0002827]; positive regulation of transcription from RNA polymerase II promoter [GO:0045944]; respiratory burst [GO:0045730]; response to bacterium [GO:0009617]; response to interferon-gamma [GO:0034341]; response to lipopolysaccharide [GO:0032496]; T cell cytokine production [GO:0002369]; T cell proliferation involved in immune response [GO:0002309]; vacuolar acidification [GO:0007035]; wound healing [GO:0042060]
89	P49281	NRAM2_HUMAN	SLC11A2 DCT1 DMT1 NRAMP2 OK/SW-cl.20	Natural resistance-associated macrophage protein 2 (NRAMP 2) (Divalent cation transporter 1) (Divalent metal transporter 1) (DMT-1) (Solute carrier family 11 member 2)	Unknown	Unknown	Substrate - transport		Cell membrane, Endosome	Yes	DISEASE: Anemia, hypochromic microcytic, with iron overload 1 (AHMIO1) [MIM:206100]: A hematologic disease characterized by abnormal hemoglobin content in the erythrocytes which are reduced in size. The disorder is due to an error of iron metabolism that results in high serum iron, massive hepatic iron deposition, and absence of sideroblasts and stainable bone marrow iron store. Despite adequate transferrin-iron complex, delivery of iron to the erythroid bone marrow is apparently insufficient for the demands of hemoglobin synthesis. {ECO:0000269|PubMed:15459009, ECO:0000269|PubMed:16160008, ECO:0000269|PubMed:16439678}. Note=The disease is caused by mutations affecting the gene represented in this entry.	activation of cysteine-type endopeptidase activity involved in apoptotic process [GO:0006919]; cadmium ion transmembrane transport [GO:0070574]; cellular iron ion homeostasis [GO:0006879]; cellular response to oxidative stress [GO:0034599]; cobalt ion transport [GO:0006824]; copper ion transport [GO:0006825]; dendrite morphogenesis [GO:0048813]; detection of oxygen [GO:0003032]; erythrocyte development [GO:0048821]; ferrous iron import [GO:0070627]; ferrous iron transport [GO:0015684]; heme biosynthetic process [GO:0006783]; lead ion transport [GO:0015692]; learning or memory [GO:0007611]; manganese ion transport [GO:0006828]; multicellular organismal iron ion homeostasis [GO:0060586]; nickel cation transport [GO:0015675]; response to hypoxia [GO:0001666]; response to iron ion [GO:0010039]; vanadium ion transport [GO:0015676]
90	Q8N543	OGFD1_HUMAN	OGFOD1 KIAA1612 TPA1	Prolyl 3-hydroxylase OGFOD1 (EC 1.14.11.-) (2-oxoglutarate and iron-dependent oxygenase domain-containing protein 1) (Termination and polyadenylation 1 homolog)	H155-D157-H218	1 Fe cation	Catalytic	1.14.11.-	Cytoplasm, Nucleus	No		cell proliferation [GO:0008283]; peptidyl-proline hydroxylation [GO:0019511]; protein hydroxylation [GO:0018126]; regulation of translational termination [GO:0006449]; stress granule assembly [GO:0034063]
91	Q6N063	OGFD2_HUMAN	OGFOD2	2-oxoglutarate and iron-dependent oxygenase domain-containing protein 2 (EC 1.14.11.-)	H235-D237-H290	1 Fe cation	Catalytic	1.14.11.-	Unknown	No		
92	Q6PK18	OGFD3_HUMAN	OGFOD3 C17orf101	2-oxoglutarate and iron-dependent oxygenase domain-containing protein 3 (EC 1.14.11.-)	H230-D232-H288	1 Fe cation	Catalytic	1.14.11.-	Unknown	Yes		
93	Q9NPF4	OSGEP_HUMAN	OSGEP GCPL1	Probable tRNA N6-adenosine threonylcarbamoyltransferase (EC 2.3.1.234) (N6-L-threonylcarbamoyladenine synthase) (t(6)A synthase) (O-sialoglycoprotein endopeptidase) (hOSGEP) (t(6)A37 threonylcarbamoyladenosine biosynthesis protein OSGEP) (tRNA threonylcarbamoyladenosine biosynthesis protein OSGEP)	H109-H113-Y130-D294	1 Divalent cation	Catalytic	2.3.1.234	Cytoplasm, Nucleus	No		tRNA threonylcarbamoyladenosine modification [GO:0002949]

94	Q32P28	P3H1_HUMAN	P3H1 GROS1 LEPRE1 PSEC0109	Prolyl 3-hydroxylase 1 (EC 1.14.11.7) (Growth suppressor 1) (Leucine- and proline-enriched proteoglycan 1) (Leprecan-1)	H587-D589-H659	1 Fe cation	Catalytic	1.14.11.7	Endoplasmic reticulum	No	DISEASE: Osteogenesis imperfecta 8 (OI8) [MIM:610915]: A form of osteogenesis imperfecta, a connective tissue disorder characterized by low bone mass, bone fragility and susceptibility to fractures after minimal trauma. Disease severity ranges from very mild forms without fractures to intrauterine fractures and perinatal lethality. Extraskeletal manifestations, which affect a variable number of patients, are dentinogenesis imperfecta, hearing loss, and blue sclerae. OI8 is characterized by disproportionate short stature, severe osteoporosis, shortening of the long bones, white sclerae, a round face and a short barrel-shaped chest. {ECO:0000269|PubMed:17277775, ECO:0000269|PubMed:19088120}. Note=The disease is caused by mutations affecting the gene represented in this entry. A splice site mutation leading to the absence of isoform 1 has been reported in 2 OI8 patients. Isoform 1 is the only form predicted to be located in the endoplasmic reticulum, which the appropriate location for the catalysis of collagen hydroxylation. These patients show indeed severely reduced COL1A1 hydroxylation (PubMed:19088120). {ECO:0000269|PubMed:19088120}.	bone development [GO:0060348]; chaperone-mediated protein folding [GO:0061077]; collagen metabolic process [GO:0032963]; negative regulation of cell proliferation [GO:0008285]; negative regulation of post-translational protein modification [GO:1901874]; protein folding [GO:0006457]; protein hydroxylation [GO:0018126]; protein stabilization [GO:0050821]; regulation of protein secretion [GO:0050708]
95	Q8IVL5	P3H2_HUMAN	P3H2 LEPREL1 MLAT4	Prolyl 3-hydroxylase 2 (EC 1.14.11.7) (Leprecan-like protein 1) (Myxoid liposarcoma-associated protein 4)	H580-D582-H652	1 Fe cation	Catalytic	1.14.11.7	Endoplasmic reticulum, Golgi apparatus	No	DISEASE: Myopia, high, with cataract and vitreoretinal degeneration (MCVD) [MIM:614292]: A disorder characterized by severe myopia with variable expressivity of cataract and vitreoretinal degeneration. Some patients manifest lens subluxation, lens instability and retinal detachment. {ECO:0000269|PubMed:21885030}. Note=The disease is caused by mutations affecting the gene represented in this entry.	collagen metabolic process [GO:0032963]; negative regulation of cell proliferation [GO:0008285]; peptidyl-proline hydroxylation [GO:0019511]
96	Q8IVL6	P3H3_HUMAN	P3H3 LEPREL2	Prolyl 3-hydroxylase 3 (EC 1.14.11.7) (Leprecan-like protein 2) (Protein B)	H584-D586-H656	1 Fe cation	Catalytic	1.14.11.7	Endoplasmic reticulum	No		collagen metabolic process [GO:0032963]; negative regulation of cell proliferation [GO:0008285]
97	P13674	P4HA1_HUMAN	P4HA1 P4HA	Prolyl 4-hydroxylase subunit alpha-1 (4-PH alpha-1) (EC 1.14.11.2) (Procollagen-proline,2-oxoglutarate-4-dioxygenase subunit alpha-1)	H429-D431-H500	1 Fe cation	Catalytic	1.14.11.2	Endoplasmic reticulum	No		collagen fibril organization [GO:0030199]; peptidyl-proline hydroxylation to 4-hydroxy-L-proline [GO:0018401]
98	O15460	P4HA2_HUMAN	P4HA2 UNQ290/PRO330	Prolyl 4-hydroxylase subunit alpha-2 (4-PH alpha-2) (EC 1.14.11.2) (Procollagen-proline,2-oxoglutarate-4-dioxygenase subunit alpha-2)	H430-D432-H501	1 Fe cation	Catalytic	1.14.11.2	Endoplasmic reticulum	No	DISEASE: Myopia 25, autosomal dominant (MYP25) [MIM:617238]: A refractive error of the eye, in which parallel rays from a distant object come to focus in front of the retina, vision being better for near objects than for far. {ECO:0000269|PubMed:25741866}. Note=The disease is caused by mutations affecting the gene represented in this entry.	
99	Q7Z4N8	P4HA3_HUMAN	P4HA3 UNQ711/PRO1374	Prolyl 4-hydroxylase subunit alpha-3 (4-PH alpha-3) (EC 1.14.11.2) (Procollagen-proline,2-oxoglutarate-4-dioxygenase subunit alpha-3)	H440-D442-H510	1 Fe cation	Catalytic	1.14.11.2	Endoplasmic reticulum	No		
100	Q9NXG6	P4HTM_HUMAN	P4HTM PH4	Transmembrane prolyl 4-hydroxylase (P4H-TM) (EC 1.14.11.-) (Hypoxia-inducible factor prolyl hydroxylase 4) (HIF-PH4) (HIF-prolyl hydroxylase 4) (HPH-4)	H328-D330-H441	1 Fe cation	Catalytic	1.14.11.-	Endoplasmic reticulum	Yes		regulation of erythrocyte differentiation [GO:0045646]
101	O14832	PAHX_HUMAN	PHYH PAHX	Phytanoyl-CoA dioxygenase, peroxisomal (EC 1.14.11.18) (Phytanic acid oxidase) (Phytanoyl-CoA alpha-hydroxylase) (PhyH)	H175-D177-H264	1 Fe cation	Catalytic	1.14.11.18	Peroxisome	No	DISEASE: Refsum disease (RD) [MIM:266500]: A rare autosomal recessive peroxisomal disorder characterized by the accumulation of the branched-chain fatty acid, phytanic acid, in blood and tissues. Cardinal clinical features are retinitis pigmentosa, peripheral neuropathy, cerebellar ataxia, and elevated protein levels in the cerebrospinal fluid (CSF). Half of all patients exhibit generalized, mild to moderate ichthyosis resembling ichthyosis vulgaris. Less constant features are nerve deafness, anosmia, skeletal abnormalities, cataracts and cardiac impairment. {ECO:0000269|PubMed:10709665, ECO:0000269|PubMed:10767344, ECO:0000269|PubMed:14974078, ECO:0000269|PubMed:9326939, ECO:0000269|PubMed:9326940}. Note=The disease is caused by mutations affecting the gene represented in this entry.	2-oxoglutarate metabolic process [GO:0006103]; fatty acid alpha-oxidation [GO:0001561]; isoprenoid metabolic process [GO:0006720]; methyl-branched fatty acid metabolic process [GO:0097089]

102	P00439	PH4H_HUMAN	PAH	Phenylalanine-4-hydroxylase (PAH) (EC 1.14.16.1) (Phe-4-monooxygenase)	H285-H290-E330	1 Fe cation	Catalytic	1.14.16.1	Unknown	No	DISEASE: Phenylketonuria (PKU) [MIM:261600]: Autosomal recessive inborn error of phenylalanine metabolism, due to severe phenylalanine hydroxylase deficiency. It is characterized by blood concentrations of phenylalanine persistently above 1200 mumol (normal concentration 100 mumol) which usually causes mental retardation (unless low phenylalanine diet is introduced early in life). They tend to have light pigmentation, rashes similar to eczema, epilepsy, extreme hyperactivity, psychotic states and an unpleasant 'mousy' odor. {ECO:0000269|PubMed:10200057, ECO:0000269|PubMed:10679941, ECO:0000269|PubMed:11180595, ECO:0000269|PubMed:11326337, ECO:0000269|PubMed:11385716, ECO:0000269|PubMed:11461196, ECO:0000269|PubMed:12501224, ECO:0000269|PubMed:1355066, ECO:0000269|PubMed:1363837, ECO:0000269|PubMed:1363838, ECO:0000269|PubMed:1671810, ECO:0000269|PubMed:1672290, ECO:0000269|PubMed:1672294, ECO:0000269|PubMed:1679030, ECO:0000269|PubMed:1709636, ECO:0000269|PubMed:18538294, ECO:0000269|PubMed:1975559, ECO:0000269|PubMed:2014802, ECO:0000269|PubMed:22513348, ECO:0000269|PubMed:22526846, ECO:0000269|PubMed:23792259, ECO:0000269|PubMed:2564729, ECO:0000269|PubMed:2615649, ECO:0000269|PubMed:2840952, ECO:0000269|PubMed:7833954, ECO:0000269|PubMed:8068076, ECO:0000269|PubMed:8406445, ECO:0000269|PubMed:8889583, ECO:0000269|PubMed:8889590, ECO:0000269|PubMed:9048935, ECO:0000269|PubMed:9101291, ECO:0000269|PubMed:9452061, ECO:0000269|PubMed:9452062, ECO:0000269|PubMed:9521426, ECO:0000269|PubMed:9600453, ECO:0000269|PubMed:9792407, ECO:0000269|PubMed:9792411, ECO:0000269|PubMed:9950317}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Non-phenylketonuria hyperphenylalaninemia (Non-PKU HPA) [MIM:261600]: Mild form of phenylalanine hydroxylase deficiency characterized by phenylalanine levels persistently below 600 mumol, which allows normal intellectual and behavioral development without treatment. Non-PKU HPA is usually caused by the combined effect of a mild hyperphenylalaninemia mutation and a severe one. {ECO:0000269|PubMed:1358789, ECO:0000269|PubMed:8088845, ECO:0000269|PubMed:8098245, ECO:0000269|PubMed:9521426, ECO:0000269|PubMed:9852673}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Hyperphenylalaninemia (HPA) [MIM:261600]: Mildest form of phenylalanine hydroxylase deficiency. {ECO:0000269|PubMed:11385716, ECO:0000269|PubMed:11935335, ECO:0000269|PubMed:12501224, ECO:0000269|PubMed:1358789, ECO:0000269|PubMed:23792259, ECO:0000269|PubMed:8088845, ECO:0000269|PubMed:8098245, ECO:0000269|PubMed:9521426, ECO:0000269|PubMed:9852673}. Note=The disease is caused by mutations affecting the gene represented in this entry.	catecholamine biosynthetic process [GO:0042423]; cellular amino acid biosynthetic process [GO:0008652]; L-phenylalanine catabolic process [GO:0006559]; neurotransmitter biosynthetic process [GO:0042136]
103	O75151	PHF2_HUMAN	PHF2 CENP-35 KIAA0662	Lysine-specific demethylase PHF2 (EC 1.14.11.-) (GRC5) (PHD finger protein 2)	H249-D251-Y321	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No		liver development [GO:0001889]; negative regulation of chromatin silencing at rDNA [GO:0061188]; protein demethylation [GO:0006482]; transcription, DNA-templated [GO:0006351]

104	Q9UPP1	PHF8_HUMAN	PHF8 KIAA1111 ZNF422	Histone lysine demethylase PHF8 (EC 1.14.11.27) (PHD finger protein 8)	H283-D285-Y293-H355	1 Fe cation	Catalytic	1.14.11.27	Nucleus	No	DISEASE: Mental retardation, X-linked, syndromic, Siderius type (MRXSSD) [MIM:300263]: A syndrome characterized by mild to borderline mental retardation with or without cleft lip/cleft palate. {ECO:0000269|PubMed:16199551, ECO:0000269|PubMed:17661819, ECO:0000269|PubMed:20101266, ECO:0000269|PubMed:20208542, ECO:0000269|PubMed:20346720, ECO:0000269|PubMed:20421419, ECO:0000269|PubMed:20548336, ECO:0000269|PubMed:20622853, ECO:0000269|PubMed:20622854}. Note=The disease is caused by mutations affecting the gene represented in this entry.	brain development [GO:0007420]; G1/S transition of mitotic cell cycle [GO:0000082]; histone H3-K27 demethylation [GO:0071557]; histone H3-K36 demethylation [GO:0070544]; histone H3-K9 demethylation [GO:0033169]; histone H4-K20 demethylation [GO:0035574]; negative regulation of chromatin silencing at rDNA [GO:0061188]; positive regulation of transcription, DNA-templated [GO:0045893]; positive regulation of transcription from RNA polymerase I promoter [GO:0045943]; transcription, DNA-templated [GO:0006351]
105	Q5SRE7	PHYD1_HUMAN	PHYHD1	Phytanoyl-CoA dioxygenase domain-containing protein 1 (EC 1.-.-.-)	H156-D158-H246	1 Fe cation	Catalytic	1.-.-.-	Unknown	No		
106	O00625	PIR_HUMAN	PIR	Pirin (EC 1.13.11.24) (Probable quercetin 2,3-dioxygenase PIR) (Probable quercetinase)	H56-H58-H101-E103	1 Fe cation	Catalytic	1.13.11.24	Cytoplasm, Nucleus	No		monocyte differentiation [GO:0030224]; regulation of transcription, DNA-templated [GO:0006355]; transcription from RNA polymerase II promoter [GO:0006366]
107	Q02809	PLOD1_HUMAN	PLOD1 LLH PLOD	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1 (EC 1.14.11.4) (Lysyl hydroxylase 1) (LH1)	H656-D658-H708	1 Fe cation	Catalytic	1.14.11.4	Endoplasmic reticulum	Yes	DISEASE: Ehlers-Danlos syndrome 6 (EDS6) [MIM:225400]: A connective tissue disorder characterized by generalized joint hypermobility, hyperextensible skin, atrophic cutaneous scars due to tissue fragility, progressive kyphoscoliosis already present at birth, ocular manifestations, arterial rupture, easy bruising, severe neonatal muscle hypotonia and delayed motor development. {ECO:0000269|PubMed:10686424, ECO:0000269|PubMed:15666309, ECO:0000269|PubMed:15854030, ECO:0000269|PubMed:15979919, ECO:0000269|PubMed:8163671, ECO:0000269|PubMed:9617436}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cellular protein modification process [GO:0006464]; epidermis development [GO:0008544]; hydroxylysine biosynthetic process [GO:0046947]; oxidation-reduction process [GO:0055114]; peptidyl-lysine hydroxylation [GO:0017185]; response to hypoxia [GO:0001666]
108	O00469	PLOD2_HUMAN	PLOD2	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2 (EC 1.14.11.4) (Lysyl hydroxylase 2) (LH2)	H666-D668-H718	1 Fe cation	Catalytic	1.14.11.4	Endoplasmic reticulum	Yes	DISEASE: Bruck syndrome 2 (BRKS2) [MIM:609220]: An autosomal recessive disease characterized by generalized osteopenia, congenital joint contractures, fragile bones with onset of fractures in infancy or early childhood, short stature, severe limb deformity, progressive scoliosis, and pterygia. It is distinguished from osteogenesis imperfecta by the absence of hearing loss and dentinogenesis imperfecta, and by the presence of clubfoot and congenital joint limitations. {ECO:0000269|PubMed:12881513, ECO:0000269|PubMed:15523624}. Note=The disease is caused by mutations affecting the gene represented in this entry. The molecular defect leading to Bruck syndrome is an aberrant cross-linking of bone collagen, due to underhydroxylation of lysine residues within the telopeptides of type I collagen, whereas the lysine residues in the triple helix are normal.; DISEASE: Note=PLOD2 mutations give rise to a broad variety of phenotypes with variable degrees of severity of bone fragility and joint contractures. Disease-associated mutations have been found in patients with autosomal recessive osteogenesis imperfecta (AR-OI) (PubMed:22689593). {ECO:0000269|PubMed:22689593}.	cellular protein modification process [GO:0006464]; hydroxylysine biosynthetic process [GO:0046947]; peptidyl-lysine hydroxylation [GO:0017185]; response to hypoxia [GO:0001666]
109	O60568	PLOD3_HUMAN	PLOD3	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 3 (EC 1.14.11.4) (Lysyl hydroxylase 3) (LH3)	H667-D669-H719	1 Fe cation	Catalytic	1.14.11.4	Endoplasmic reticulum	Yes	DISEASE: Lysyl hydroxylase 3 deficiency (LH3 deficiency) [MIM:612394]: Connective tissue disorder. The syndrome is characterized by congenital malformations severely affecting many tissues and organs and revealing features of several collagen disorders, most of them involving COL2A1 (type II collagen). The findings suggest that the failure of lysyl hydroxylation and hydroxylysyl carbohydrate addition, which affects many collagens, is the molecular basis of this syndrome. {ECO:0000269|PubMed:18834968}. Note=The disease is caused by mutations affecting the gene represented in this entry.	basement membrane assembly [GO:0070831]; cellular response to hormone stimulus [GO:0032870]; collagen fibril organization [GO:0030199]; endothelial cell morphogenesis [GO:0001886]; epidermis morphogenesis [GO:0048730]; hydroxylysine biosynthetic process [GO:0046947]; in utero embryonic development [GO:0001701]; lung morphogenesis [GO:0060425]; neural tube development [GO:0021915]; peptidyl-lysine hydroxylation [GO:0017185]; protein localization [GO:0008104]; protein O-linked glycosylation [GO:0006493]; vasodilation [GO:0042311]

110	P62136	PP1A_HUMAN	PPP1CA PPP1A	Serine/threonine-protein phosphatase PP1-alpha catalytic subunit (PP-1A) (EC 3.1.3.16)	D64-H66-D92	1 Fe cation	Catalytic	3.1.3.16	Cytoplasm, Nucleus	No		beta-catenin destruction complex disassembly [GO:1904886]; branching morphogenesis of an epithelial tube [GO:0048754]; cell cycle [GO:0007049]; cell division [GO:0051301]; circadian regulation of gene expression [GO:0032922]; dephosphorylation [GO:0016311]; entrainment of circadian clock by photoperiod [GO:0043153]; glycogen metabolic process [GO:0005977]; lung development [GO:0030324]; negative regulation of protein binding [GO:0032091]; positive regulation of extrinsic apoptotic signaling pathway in absence of ligand [GO:2001241]; protein dephosphorylation [GO:0006470]; regulation of canonical Wnt signaling pathway [GO:0060828]; regulation of circadian rhythm [GO:0042752]; regulation of glycogen biosynthetic process [GO:0005979]; regulation of glycogen catabolic process [GO:0005981]; regulation of translational initiation by eIF2 alpha dephosphorylation [GO:0036496]
111	Q08209	PP2BA_HUMAN	PPP3CA CALNA CNA	Serine/threonine-protein phosphatase 2B catalytic subunit alpha isoform (EC 3.1.3.16) (CAM-PRP catalytic subunit) (Calmodulin-dependent calcineurin A subunit alpha isoform)	D90-H92-D118	1 Fe cation	Catalytic	3.1.3.16	Cell membrane, Nucleus	Yes		calcineurin-NFAT signaling cascade [GO:0033173]; calcium ion transport [GO:0006816]; cardiac muscle hypertrophy in response to stress [GO:0014898]; cellular response to drug [GO:0035690]; cellular response to glucose stimulus [GO:0071333]; dephosphorylation [GO:0016311]; excitatory postsynaptic potential [GO:0060079]; Fc-epsilon receptor signaling pathway [GO:0038095]; G1/S transition of mitotic cell cycle [GO:0000082]; modulation of synaptic transmission [GO:0050804]; multicellular organismal response to stress [GO:0033555]; negative regulation of chromatin binding [GO:0035562]; negative regulation of dendrite morphogenesis [GO:0050774]; negative regulation of insulin secretion [GO:0046676]; negative regulation of production of miRNAs involved in gene silencing by miRNA [GO:1903799]; positive regulation of cardiac muscle hypertrophy in response to stress [GO:1903244]; positive regulation of connective tissue replacement [GO:1905205]; positive regulation of NFAT protein import into nucleus [GO:0051533]; positive regulation of sequence-specific DNA binding transcription factor activity [GO:0051091]; positive regulation of transcription from RNA polymerase II promoter [GO:0045944]; protein dephosphorylation [GO:0006470]; protein import into nucleus [GO:0006606]; response to amphetamine [GO:0001975]; response to calcium ion [GO:0051592]; skeletal muscle fiber development [GO:0048741]; T cell activation [GO:0042110]; transition between fast and slow fiber [GO:0014883]; Wnt signaling pathway, calcium modulating pathway [GO:0007223]
112	P16298	PP2BB_HUMAN	PPP3CB CALNA2 CALNB CNA2	Serine/threonine-protein phosphatase 2B catalytic subunit beta isoform (EC 3.1.3.16) (CAM-PRP catalytic subunit) (Calmodulin-dependent calcineurin A subunit beta isoform)	D99-H101-D127	1 Fe cation	Catalytic	3.1.3.16	Unknown	No		axon extension [GO:0048675]; calcineurin-NFAT signaling cascade [GO:0033173]; calcium ion regulated exocytosis [GO:0017156]; cellular response to drug [GO:0035690]; dephosphorylation [GO:0016311]; Fc-epsilon receptor signaling pathway [GO:0038095]; heart development [GO:0007507]; learning [GO:0007612]; locomotion involved in locomotory behavior [GO:0031987]; lymphangiogenesis [GO:0001946]; memory [GO:0007613]; negative regulation of T cell mediated cytotoxicity [GO:0001915]; positive regulation of insulin secretion involved in cellular response to glucose stimulus [GO:0035774]; positive regulation of NFAT protein import into nucleus [GO:0051533]; positive regulation of transcription, DNA-templated [GO:0045893]; positive regulation of transcription from RNA polymerase II promoter [GO:0045944]; protein dephosphorylation [GO:0006470]; protein phosphorylation [GO:0006468]; regulation of insulin secretion [GO:0050796]; regulation of synaptic plasticity [GO:0048167]; response to cytokine [GO:0034097]; signal transduction [GO:0007165]; social behavior [GO:0035176]; T cell activation [GO:0042110]; T cell differentiation [GO:0030217]; T cell homeostasis [GO:0043029]; T cell proliferation [GO:0042098]; Wnt signaling pathway, calcium modulating pathway [GO:0007223]
113	P48454	PP2BC_HUMAN	PPP3CC CALNA3 CNA3	Serine/threonine-protein phosphatase 2B catalytic subunit gamma isoform (EC 3.1.3.16) (CAM-PRP catalytic subunit) (Calcineurin, testis-specific catalytic subunit) (Calmodulin-dependent calcineurin A subunit gamma isoform)	D86-H88-D114	1 Fe cation	Catalytic	3.1.3.16	Unknown	No		brain development [GO:0007420]; positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway [GO:1900740]

114	P13686	PPA5_HUMAN	ACP5	Tartrate-resistant acid phosphatase type 5 (TR-AP) (EC 3.1.3.2) (Tartrate-resistant acid ATPase) (TrATPase) (Type 5 acid phosphatase)	D33-D71-Y74-H242; D71-N110-H205-H240	2 Fe cations	Catalytic	3.1.3.2	Unknown	No	DISEASE: Spondyloenchondrodysplasia with immune dysregulation (SPENCDI) [MIM:607944]: A disease characterized by vertebral and metaphyseal dysplasia, spasticity with cerebral calcifications, and strong predisposition to autoimmune diseases. The skeletal dysplasia is characterized by radiolucent and irregular spondylar and metaphyseal lesions that represent islands of chondroid tissue within bone. {ECO:0000269|PubMed:21217752, ECO:0000269|PubMed:21217755}. Note=The disease is caused by mutations affecting the gene represented in this entry. ACP5 inactivating mutations result in a functional excess of phosphorylated osteopontin causing deregulation of osteopontin signaling and consequential autoimmune disease.	riboflavin metabolic process [GO:0006771]
115	Q7KZA3	Q7KZA3_HUMAN	DKFZp686P18130	Ferrochelatase	Unknown	1 Fe cation	Substrate - biosinthesis	4.99.1.1	Unknown	No		ferrochelatase activity
116	Q9H6W3	RIOX1_HUMAN	RIOX1 C14orf169 MAPJD NO66	Ribosomal oxygenase 1 (60S ribosomal protein L8 histidine hydroxylase) (Bifunctional lysine-specific demethylase and histidyl-hydroxylase NO66) (EC 1.14.11.-) (EC 1.14.11.27) (Histone lysine demethylase NO66) (Myc-associated protein with JmjC domain) (Nucleolar protein 66) (hsNO66) (Ribosomal oxygenase NO66) (ROX)	H340-D342-H405	1 Fe cation	Catalytic	1.14.11.-; 1.14.11.27	Nucleus	No		chromatin remodeling [GO:0006338]; histone H3-K36 demethylation [GO:0070544]; histone H3-K4 demethylation [GO:0034720]; negative regulation of osteoblast differentiation [GO:0045668]; negative regulation of transcription, DNA-templated [GO:0045892]; peptidyl-arginine hydroxylation [GO:0030961]; transcription, DNA-templated [GO:0006351]
117	Q8IUF8	RIOX2_HUMAN	RIOX2 MDIG MINA MINA53 NO52	Ribosomal oxygenase 2 (60S ribosomal protein L27a histidine hydroxylase) (Bifunctional lysine-specific demethylase and histidyl-hydroxylase MINA) (EC 1.14.11.-) (Histone lysine demethylase MINA) (MYC-induced nuclear antigen) (Mineral dust-induced gene protein) (Nucleolar protein 52) (Ribosomal oxygenase MINA) (ROX)	H179-D181-H240	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No		chromatin remodeling [GO:0006338]; negative regulation of transcription, DNA-templated [GO:0045892]; peptidyl-arginine hydroxylation [GO:0030961]; ribosome biogenesis [GO:0042254]; transcription, DNA-templated [GO:0006351]
118	P31350	RIR2_HUMAN	RRM2 RR2	Ribonucleoside-diphosphate reductase subunit M2 (EC 1.17.4.1) (Ribonucleotide reductase small chain) (Ribonucleotide reductase small subunit)	D138-E169-H172; E169-E232-E266-H269	2 Fe cations	Catalytic	1.17.4.1	Cytoplasm	No		deoxyribonucleotide biosynthetic process [GO:0009263]; DNA replication [GO:0006260]; G1/S transition of mitotic cell cycle [GO:0000082]; nucleobase-containing small molecule interconversion [GO:0015949]; protein heterotetramerization [GO:0051290]; regulation of transcription involved in G1/S transition of mitotic cell cycle [GO:0000083]
119	Q7LG56	RIR2B_HUMAN	RRM2B P53R2	Ribonucleoside-diphosphate reductase subunit M2 B (EC 1.17.4.1) (TP53-inducible ribonucleotide reductase M2 B) (p53-inducible ribonucleotide reductase small subunit 2-like protein) (p53R2)	D100-E131-H134; E131-E194-E228-H231	2 Fe cations	Catalytic	1.17.4.1	Cytoplasm, Nucleus	No	DISEASE: Mitochondrial DNA depletion syndrome 8A (MTDPS8A) [MIM:612075]: A disorder due to mitochondrial dysfunction characterized by various combinations of neonatal hypotonia, neurological deterioration, respiratory distress, lactic acidosis, and renal tubulopathy. {ECO:0000269|PubMed:17486094, ECO:0000269|PubMed:18504129}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Mitochondrial DNA depletion syndrome 8B (MTDPS8B) [MIM:612075]: A disease due to mitochondrial dysfunction and characterized by ophthalmoplegia, ptosis, gastrointestinal dysmotility, cachexia, peripheral neuropathy. {ECO:0000269|PubMed:19667227}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Progressive external ophthalmoplegia with mitochondrial DNA deletions, autosomal dominant, 5 (PEOA5) [MIM:613077]: A disorder characterized by progressive weakness of ocular muscles and levator muscle of the upper eyelid. In a minority of cases, it is associated with skeletal myopathy, which predominantly involves axial or proximal muscles and which causes abnormal fatigability and even permanent muscle weakness. Ragged-red fibers and atrophy are found on muscle biopsy. A large proportion of chronic ophthalmoplegias are associated with other symptoms, leading to a multisystemic pattern of this disease. Additional symptoms are variable, and may include cataracts, hearing loss, sensory axonal neuropathy, ataxia, depression, hypogonadism, and parkinsonism. {ECO:0000269|PubMed:19664747}. Note=The disease is caused by mutations affecting the gene represented in this entry.	deoxyribonucleoside triphosphate metabolic process [GO:0009200]; deoxyribonucleotide biosynthetic process [GO:0009263]; DNA repair [GO:0006281]; kidney development [GO:0001822]; mitochondrial DNA replication [GO:0006264]; negative regulation of intrinsic apoptotic signaling pathway by p53 class mediator [GO:1902254]; nucleobase-containing small molecule interconversion [GO:0015949]; renal system process [GO:0003014]; response to amine [GO:0014075]; response to oxidative stress [GO:0006979]

120	Q96AT9	RPE_HUMAN	RPE HUSSY-17	Ribulose-phosphate 3-epimerase (EC 5.1.3.1) (Ribulose-5-phosphate-3-epimerase)	H35-D37-H70-D175	1 Divalent cation	Catalytic - no redox	5.1.3.1	Unknown	No		carbohydrate metabolic process [GO:0005975]; cellular carbohydrate metabolic process [GO:0044262]; pentose catabolic process [GO:0019323]; pentose-phosphate shunt [GO:0006098]; pentose-phosphate shunt, non-oxidative branch [GO:0009052]
121	Q16518	RPE65_HUMAN	RPE65	Retinoid isomerohydrolase (EC 3.1.1.64) (All-trans-retinyl-palmitate hydrolase) (Retinal pigment epithelium-specific 65 kDa protein) (Retinol isomerase)	H180-H241-H313-H527	1 Fe cation	Catalytic	3.1.1.64	Cytoplasm, Cell membrane	Yes	DISEASE: Leber congenital amaurosis 2 (LCA2) [MIM:204100]: A severe dystrophy of the retina, typically becoming evident in the first years of life. Visual function is usually poor and often accompanied by nystagmus, sluggish or near-absent pupillary responses, photophobia, high hyperopia and keratoconus. {ECO:0000269|PubMed:10090910, ECO:0000269|PubMed:10766140, ECO:0000269|PubMed:11462243, ECO:0000269|PubMed:14611946, ECO:0000269|PubMed:14962443, ECO:0000269|PubMed:15024725, ECO:0000269|PubMed:16205573, ECO:0000269|PubMed:17297704, ECO:0000269|PubMed:17724218, ECO:0000269|PubMed:18682808, ECO:0000269|PubMed:9326941, ECO:0000269|PubMed:9801879}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Retinitis pigmentosa 20 (RP20) [MIM:613794]: A retinal dystrophy belonging to the group of pigmentary retinopathies. Retinitis pigmentosa is characterized by retinal pigment deposits visible on fundus examination and primary loss of rod photoreceptor cells followed by secondary loss of cone photoreceptors. Patients typically have night vision blindness and loss of midperipheral visual field. As their condition progresses, they lose their far peripheral visual field and eventually central vision as well. {ECO:0000269|PubMed:11095629, ECO:0000269|PubMed:12960219, ECO:0000269|PubMed:15557452, ECO:0000269|PubMed:22334370, ECO:0000269|PubMed:23878505, ECO:0000269|PubMed:9501220}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Note=Defects in RPE65 are a cause of autosomal dominant retinitis pigmentosa with choroidal involvement (PubMed:21654732). Affected individuals show reduction of central vision, constriction of visual fields, night blindness and chorioretinal atrophy. {ECO:0000269|PubMed:21654732}.	cellular response to electrical stimulus [GO:0071257]; circadian rhythm [GO:0007623]; detection of light stimulus involved in visual perception [GO:0050908]; insulin receptor signaling pathway [GO:0008286]; neural retina development [GO:0003407]; regulation of rhodopsin gene expression [GO:0007468]; retina homeostasis [GO:0001895]; retinal metabolic process [GO:0042574]; retina morphogenesis in camera-type eye [GO:0060042]; retinoid metabolic process [GO:0001523]; visual perception [GO:0007601]; vitamin A metabolic process [GO:0006776]
122	Q2QD12	RPEL1_HUMAN	RPEL1	Ribulose-phosphate 3-epimerase-like protein 1 (EC 5.1.3.1) (Ribulose-5-phosphate-3-epimerase-like protein 1)	H35-D37-H70-D175	1 Divalent cation	Catalytic - no redox	5.1.3.1	Unknown	No		cellular carbohydrate metabolic process [GO:0044262]; pentose catabolic process [GO:0019323]; pentose-phosphate shunt, non-oxidative branch [GO:0009052]
123	Q9NP59	S40A1_HUMAN	SLC40A1 FPN1 IREG1 SLC11A3 MSTP079	Solute carrier family 40 member 1 (Ferroportin-1) (Iron-regulated transporter 1)	Unknown	Unknown	Substrate - transport		Cell membrane	Yes	DISEASE: Hemochromatosis 4 (HFE4) [MIM:606069]: A disorder of iron metabolism characterized by iron overload. Excess iron is deposited in a variety of organs leading to their failure, and resulting in serious illnesses including cirrhosis, hepatomas, diabetes, cardiomyopathy, arthritis, and hypogonadotropic hypogonadism. Severe effects of the disease usually do not appear until after decades of progressive iron loading. {ECO:0000269|PubMed:10747949, ECO:0000269|PubMed:11431687, ECO:0000269|PubMed:11518736, ECO:0000269|PubMed:12091366, ECO:0000269|PubMed:12091367, ECO:0000269|PubMed:12123233, ECO:0000269|PubMed:12406098, ECO:0000269|PubMed:12730114, ECO:0000269|PubMed:12857562, ECO:0000269|PubMed:12865285, ECO:0000269|PubMed:15338274, ECO:0000269|PubMed:15466004, ECO:0000269|PubMed:16351644}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cellular iron ion homeostasis [GO:0006879]; endothelium development [GO:0003158]; ferrous iron export across plasma membrane [GO:1903988]; iron ion transmembrane transport [GO:0034755]; lymphocyte homeostasis [GO:0002260]; multicellular organismal iron ion homeostasis [GO:0060586]; negative regulation of apoptotic process [GO:0043066]; positive regulation of transcription by RNA polymerase II [GO:0045944]; regulation of transcription from RNA polymerase II promoter in response to iron [GO:0034395]; spleen trabecula formation [GO:0060345]
124	O75845	SC5D_HUMAN	SC5D SC5DL	Lathosterol oxidase (EC 1.14.19.20) (C-5 sterol desaturase) (Delta(7)-sterol 5-desaturase) (Delta(7)-sterol C5(6)-desaturase) (Lathosterol 5-desaturase) (Sterol-C5-desaturase)	H138-H142-H151-H155; H209-H228-H232-H233	2 Fe cations	Catalytic	1.14.19.20	Endoplasmic reticulum	Yes	DISEASE: Lathosterolosis (LATHST) [MIM:607330]: Autosomal recessive disorder characterized by a complex phenotype, including multiple congenital anomalies, mental retardation, and liver disease. {ECO:0000269|PubMed:12189593, ECO:0000269|PubMed:12812989}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cholesterol biosynthetic process via desmosterol [GO:0033489]; cholesterol biosynthetic process via lathosterol [GO:0033490]; lipid metabolic process [GO:0006629]

125	Q86SK9	SCD5_HUMAN	SCD5 ACOD4 SCD2 SCD4	Stearoyl-CoA desaturase 5 (EC 1.14.19.1) (Acyl-CoA-desaturase 4) (HSCD5) (Stearoyl-CoA 9-desaturase) (Stearoyl-CoA desaturase 2)	H94-H99-H131-H134; H135-H243-H272-H276	2 Fe cations	Catalytic	1.14.19.1	Endoplasmic reticulum	Yes		long-chain fatty-acyl-CoA biosynthetic process [GO:0035338]; unsaturated fatty acid biosynthetic process [GO:0006636]
126	Q8NFU7	TET1_HUMAN	TET1 CXXC6 KIAA1676 LCX	Methylcytosine dioxygenase TET1 (EC 1.14.11.n2) (CXXC-type zinc finger protein 6) (Leukemia-associated protein with a CXXC domain) (Ten-eleven translocation 1 gene protein)	H1672-D1674-H2028	1 Fe cation	Catalytic	1.14.11.n2	Nucleus	No	DISEASE: Note=A chromosomal aberration involving TET1 may be a cause of acute leukemias (PubMed:12646957). Translocation t(10;11)(q22;q23) with KMT2A/MLL1. This is a rare chromosomal translocation 5' KMT2A/MLL1-TET1 3' (PubMed:12124344, PubMed:12646957). {ECO:0000269|PubMed:12124344, ECO:0000269|PubMed:12646957}.	covalent chromatin modification [GO:0016569]; DNA demethylation [GO:0080111]; inner cell mass cell differentiation [GO:0001826]; negative regulation of methylation-dependent chromatin silencing [GO:0090310]; positive regulation of cell proliferation [GO:0008284]; positive regulation of histone methylation [GO:0031062]; positive regulation of transcription from RNA polymerase II promoter [GO:0045944]; protein O-linked glycosylation [GO:0006493]; stem cell population maintenance [GO:0019827]; transcription, DNA-templated [GO:0006351]
127	Q6N021	TET2_HUMAN	TET2 KIAA1546 Nbla00191	Methylcytosine dioxygenase TET2 (EC 1.14.11.n2)	H1382-D1384-H1881	1 Fe cation	Catalytic	1.14.11.n2	Unknown	No	DISEASE: Note=TET2 is frequently mutated in myeloproliferative disorders (MPD). These constitute a heterogeneous group of disorders, also known as myeloproliferative diseases or myeloproliferative neoplasms (MPN), characterized by cellular proliferation of one or more hematologic cell lines in the peripheral blood, distinct from acute leukemia. Included diseases are: essential thrombocythemia, polycythemia vera, primary myelofibrosis (chronic idiopathic myelofibrosis). Bone marrow samples from patients display uniformly low levels of hmC in genomic DNA compared to bone marrow samples from healthy controls as well as hypomethylation relative to controls at the majority of differentially methylated CpG sites.; DISEASE: Polycythemia vera (PV) [MIM:263300]: A myeloproliferative disorder characterized by abnormal proliferation of all hematopoietic bone marrow elements, erythroid hyperplasia, an absolute increase in total blood volume, but also by myeloid leukocytosis, thrombocytosis and splenomegaly. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Note=TET2 is frequently mutated in systemic mastocytosis; also known as systemic mast cell disease. A condition with features in common with myeloproliferative diseases. It is a clonal disorder of the mast cell and its precursor cells. The clinical symptoms and signs of systemic mastocytosis are due to accumulation of clonally derived mast cells in different tissues, including bone marrow, skin, the gastrointestinal tract, the liver, and the spleen.; DISEASE: Myelodysplastic syndrome (MDS) [MIM:614286]: A heterogeneous group of closely related clonal hematopoietic disorders. All are characterized by a hypercellular or hypocellular bone marrow with impaired morphology and maturation, dysplasia of the myeloid, megakaryocytic and/or erythroid lineages, and peripheral blood cytopenias resulting from ineffective blood cell production. Included diseases are: refractory anemia (RA), refractory anemia with ringed sideroblasts (RARS), refractory anemia with excess blasts (RAEB), refractory cytopenia with multilineage dysplasia and ringed sideroblasts (RCMD-RS); chronic myelomonocytic leukemia (CMML) is a myelodysplastic/myeloproliferative disease. MDS is considered a premalignant condition in a subgroup of patients that often progresses to acute myeloid leukemia (AML). {ECO:0000269|PubMed:19372255, ECO:0000269|PubMed:19483684, ECO:0000269|PubMed:21057493}. Note=The disease is caused by mutations affecting the gene represented in this entry. Bone marrow samples from patients display uniformly low levels of hmC in genomic DNA compared to bone marrow samples from healthy controls as well as hypomethylation relative to controls at the majority of differentially methylated CpG sites.	5-methylcytosine catabolic process [GO:0006211]; cell cycle [GO:0007049]; cytosine metabolic process [GO:0019858]; DNA demethylation [GO:0080111]; hematopoietic stem cell homeostasis [GO:0061484]; hemoglobin metabolic process [GO:0020027]; histone H3-K4 trimethylation [GO:0080182]; kidney development [GO:0001822]; liver morphogenesis [GO:0072576]; myeloid cell differentiation [GO:0030099]; myeloid progenitor cell differentiation [GO:0002318]; positive regulation of transcription from RNA polymerase II promoter [GO:0045944]; post-embryonic development [GO:0009791]; protein O-linked glycosylation [GO:0006493]; response to organic cyclic compound [GO:0014070]; spleen development [GO:0048536]

128	O43151	TET3_HUMAN	TET3 KIAA0401	Methylcytosine dioxygenase TET3 (EC 1.14.11.n2)	H942-D944-H1538	1 Fe cation	Catalytic	1.14.11.n2	Cytoplasm, Nucleus	No		DNA demethylation [GO:0080111]; DNA demethylation of male pronucleus [GO:0044727]; histone H3-K4 trimethylation [GO:0080182]; positive regulation of transcription from RNA polymerase II promoter [GO:0045944]; protein O-linked glycosylation [GO:0006493]
129	Q9NVH6	TMLH_HUMAN	TMLHE TMLH	Trimethyllysine dioxygenase, mitochondrial (EC 1.14.11.8) (Epsilon-trimethyllysine 2-oxoglutarate dioxygenase) (Epsilon-trimethyllysine hydroxylase) (TML hydroxylase) (TML-alpha-ketoglutarate dioxygenase) (TML dioxygenase) (TMLD)	H242-D244-H389	1 Fe cation	Catalytic	1.14.11.8	Mitochondrion	No	DISEASE: Autism, X-linked 6 (AUTSX6) [MIM:300872]: A form of autism, a complex multifactorial, pervasive developmental disorder characterized by impairments in reciprocal social interaction and communication, restricted and stereotyped patterns of interests and activities, and the presence of developmental abnormalities by 3 years of age. Most individuals with autism also manifest moderate mental retardation. AUTSX6 patients may respond favorably to carnitine supplementation. {ECO:0000269|PubMed:21865298, ECO:0000269|PubMed:23092983}. Note=The disease is caused by mutations affecting the gene represented in this entry.	carnitine biosynthetic process [GO:0045329]; negative regulation of oxidoreductase activity [GO:0051354]
130	Q0P6H9	TMM62_HUMAN	TMEM62	Transmembrane protein 62	D63-H65-D99	1 Fe cation	Catalytic		Unknown	Yes		
131	Q6ZT21	TMPPE_HUMAN	TMPPE	Transmembrane protein with metallophosphoesterase domain (EC 3.1.-.-)	D214-H216-D246-H393; N277-H369-H391	2 Divalent cations	Catalytic	3.1.-.-	Unknown	Yes		
132	P17752	TPH1_HUMAN	TPH1 TPH TPRH TRPH	Tryptophan 5-hydroxylase 1 (EC 1.14.16.4) (Tryptophan 5-monooxygenase 1)	H272-H277-E317	1 Fe cation	Catalytic	1.14.16.4	Unknown	No		aromatic amino acid family metabolic process [GO:0009072]; bone remodeling [GO:0046849]; circadian rhythm [GO:0007623]; indolalkylamine biosynthetic process [GO:0046219]; mammary gland alveolus development [GO:0060749]; negative regulation of ossification [GO:0030279]; positive regulation of fat cell differentiation [GO:0045600]; response to immobilization stress [GO:0035902]; serotonin biosynthetic process [GO:0042427]
133	Q8IWU9	TPH2_HUMAN	TPH2 NTPH	Tryptophan 5-hydroxylase 2 (EC 1.14.16.4) (Neuronal tryptophan hydroxylase) (Tryptophan 5-monooxygenase 2)	H318-H323-E363	1 Fe cation	Catalytic	1.14.16.4	Unknown	No	DISEASE: Major depressive disorder (MDD) [MIM:608516]: A common psychiatric disorder. It is a complex trait characterized by one or more major depressive episodes without a history of manic, mixed, or hypomanic episodes. A major depressive episode is characterized by at least 2 weeks during which there is a new onset or clear worsening of either depressed mood or loss of interest or pleasure in nearly all activities. Four additional symptoms must also be present including changes in appetite, weight, sleep, and psychomotor activity; decreased energy; feelings of worthlessness or guilt; difficulty thinking, concentrating, or making decisions; or recurrent thoughts of death or suicidal ideation, plans, or attempts. The episode must be accompanied by distress or impairment in social, occupational, or other important areas of functioning. {ECO:0000269|PubMed:15629698}. Note=Disease susceptibility is associated with variations affecting the gene represented in this entry.; DISEASE: Attention deficit-hyperactivity disorder 7 (ADHD7) [MIM:613003]: A neurobehavioral developmental disorder primarily characterized by the coexistence of attentional problems and hyperactivity, with each behavior occurring infrequently alone. {ECO:0000269|PubMed:18347598}. Note=Disease susceptibility is associated with variations affecting the gene represented in this entry. Naturally occurring variants of TPH2 with impaired enzyme activity could cause deficiency of serotonin production and result in an increased risk of developing behavioral disorders.	aromatic amino acid family metabolic process [GO:0009072]; cellular response to lithium ion [GO:0071285]; circadian rhythm [GO:0007623]; indolalkylamine biosynthetic process [GO:0046219]; response to activity [GO:0014823]; response to calcium ion [GO:0051592]; response to estrogen [GO:0043627]; response to glucocorticoid [GO:0051384]; response to nutrient levels [GO:0031667]; serotonin biosynthetic process [GO:0042427]
134	P02787	TRFE_HUMAN	TF PRO1400	Serotransferrin (Transferrin) (Beta-1 metal-binding globulin) (Siderophilin)	D82-Y114-Y207-H268; D411-D445-Y536-H604	2 Fe cations	Substrate - transport		Extracellular space	No	DISEASE: Atransferrinemia (ATRAF) [MIM:209300]: A rare autosomal recessive disorder characterized by abnormal synthesis of transferrin leading to iron overload and microcytic hypochromic anemia. {ECO:0000269|PubMed:11110675, ECO:0000269|PubMed:15466165}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cellular iron ion homeostasis [GO:0006879]; cellular response to iron ion [GO:0071281]; ferrous iron import across plasma membrane [GO:0098707]; iron ion homeostasis [GO:0055072]; membrane organization [GO:0061024]; platelet degranulation [GO:0002576]; positive regulation of receptor-mediated endocytosis [GO:0048260]; regulation of protein stability [GO:0031647]; retina homeostasis [GO:0001895]; transferrin transport [GO:0033572]

135	P02788	TRFL_HUMAN	LTF GIG12 LF	Lactotransferrin (Lactoferrin) (EC 3.4.21.-) (Growth-inhibiting protein 12) (Talalactoferrin) [Cleaved into: Lactoferricin-H (Lfcin-H); Kaliocin-1; Lactoferroxin-A; Lactoferroxin-B; Lactoferroxin-C]	D79-Y111-Y211-H272; D414-Y454-Y547-H616	2 Fe cations	Substrate - transport	3.4.21.-	Cytoplasm, Extracellular space	No		antibacterial humoral response [GO:0019731]; antifungal humoral response [GO:0019732]; antimicrobial humoral response [GO:0019730]; bone morphogenesis [GO:0060349]; cellular protein metabolic process [GO:0044267]; humoral immune response [GO:0006959]; innate immune response in mucosa [GO:0002227]; ion transport [GO:0006811]; iron assimilation by chelation and transport [GO:0033214]; negative regulation by host of viral process [GO:0044793]; negative regulation of apoptotic process [GO:0043066]; negative regulation of ATPase activity [GO:0032780]; negative regulation of cysteine-type endopeptidase activity [GO:2000117]; negative regulation of lipopolysaccharide-mediated signaling pathway [GO:0031665]; negative regulation of osteoclast development [GO:2001205]; negative regulation of single-species biofilm formation in or on host organism [GO:1900229]; negative regulation of tumor necrosis factor (ligand) superfamily member 11 production [GO:2000308]; negative regulation of viral genome replication [GO:0045071]; negative regulation of viral process [GO:0048525]; neutrophil degranulation [GO:0043312]; ossification [GO:0001503]; positive regulation of bone mineralization involved in bone maturation [GO:1900159]; positive regulation of chondrocyte proliferation [GO:1902732]; positive regulation of I-kappaB kinase/NF-kappaB signaling [GO:0043123]; positive regulation of NF-kappaB transcription factor activity [GO:0051092]; positive regulation of osteoblast differentiation [GO:0045669]; positive regulation of osteoblast proliferation [GO:0033690]; positive regulation of protein serine/threonine kinase activity [GO:0071902]; positive regulation of toll-like receptor 4 signaling pathway [GO:0034145]; regulation of cytokine production [GO:0001817]; regulation of tumor necrosis factor production [GO:0032680]; retina homeostasis [GO:0001895]; transcription, DNA-templated [GO:0006351]
136	P08582	TRFM_HUMAN	MELTF MAP97 MFI2	Melanotransferrin (Melanoma-associated antigen p97) (CD antigen CD228)	D78-Y107-Y210-H279; Y451-Y556-H625	2 Fe cations	Substrate - transport		Cell membrane	Yes		C-terminal protein lipidation [GO:0006501]; iron ion homeostasis [GO:0055072]; iron ion import [GO:0097286]; negative regulation of substrate adhesion-dependent cell spreading [GO:1900025]; positive regulation of extracellular matrix disassembly [GO:0090091]; positive regulation of plasminogen activation [GO:0010756]

137	P07101	TY3H_HUMAN	TH TYH	Tyrosine 3-monooxygenase (EC 1.14.16.2) (Tyrosine 3-hydroxylase) (TH)	H361-H366-E406	1 Fe cation	Catalytic	1.14.16.2	Unknown	No	DISEASE: Segawa syndrome autosomal recessive (ARSEGS) [MIM:605407]: A form of DOPA-responsive dystonia presenting in infancy or early childhood. Dystonia is defined by the presence of sustained involuntary muscle contractions, often leading to abnormal postures. Some cases present with parkinsonian symptoms in infancy. Unlike all other forms of dystonia, it is an eminently treatable condition, due to a favorable response to L-DOPA. {ECO:0000269|PubMed:10585338, ECO:0000269|PubMed:11196107, ECO:0000269|PubMed:11246459, ECO:0000269|PubMed:15505183, ECO:0000269|PubMed:15747353, ECO:0000269|PubMed:16049992, ECO:0000269|PubMed:17696123, ECO:0000269|PubMed:18058633, ECO:0000269|PubMed:18554280, ECO:0000269|PubMed:19491146, ECO:0000269|PubMed:20056467, ECO:0000269|PubMed:20430833, ECO:0000269|PubMed:21940685, ECO:0000269|PubMed:22264700, ECO:0000269|PubMed:22815559, ECO:0000269|PubMed:23762320, ECO:0000269|PubMed:23939262, ECO:0000269|PubMed:24753243, ECO:0000269|PubMed:7814018, ECO:0000269|PubMed:8528210, ECO:0000269|PubMed:8817341, ECO:0000269|PubMed:9613851, ECO:0000269|PubMed:9703425}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Note=May play a role in the pathogenesis of Parkinson disease (PD). A genome-wide copy number variation analysis has identified a 34 kilobase deletion over the TH gene in a PD patient but not in any controls. {ECO:0000269|PubMed:20809526}.	aminergic neurotransmitter loading into synaptic vesicle [GO:0015842]; anatomical structure morphogenesis [GO:0009653]; animal organ morphogenesis [GO:0009887]; catecholamine biosynthetic process [GO:0042423]; cellular response to drug [GO:0035690]; cellular response to glucose stimulus [GO:0071333]; cellular response to growth factor stimulus [GO:0071363]; cellular response to manganese ion [GO:0071287]; cellular response to nicotine [GO:0071316]; cerebral cortex development [GO:0021987]; circadian sleep/wake cycle [GO:0042745]; dopamine biosynthetic process [GO:0042416]; dopamine biosynthetic process from tyrosine [GO:0006585]; eating behavior [GO:0042755]; embryonic camera-type eye morphogenesis [GO:0048596]; epinephrine biosynthetic process [GO:0042418]; eye photoreceptor cell development [GO:0042462]; fatty acid metabolic process [GO:0006631]; glycoside metabolic process [GO:0016137]; heart development [GO:0007507]; heart morphogenesis [GO:0003007]; isoquinoline alkaloid metabolic process [GO:0033076]; learning [GO:0007612]; locomotory behavior [GO:0007626]; mating behavior [GO:0007617]; memory [GO:0007613]; multicellular organism aging [GO:0010259]; neurotransmitter biosynthetic process [GO:0042136]; norepinephrine biosynthetic process [GO:0042421]; phthalate metabolic process [GO:0018963]; phytoalexin metabolic process [GO:0052314]; pigmentation [GO:0043473]; regulation of heart contraction [GO:0008016]; response to activity [GO:0014823]; response to amphetamine [GO:0001975]; response to corticosterone [GO:0051412]; response to electrical stimulus [GO:0051602]; response to estradiol [GO:0032355]; response to ethanol [GO:0045471]; response to ether [GO:0045472]; response to herbicide [GO:0009635]; response to hypoxia [GO:0001666]; response to immobilization stress [GO:0035902]; response to isolation stress [GO:0035900]; response to light stimulus [GO:0009416]; response to lipopolysaccharide [GO:0032496]; response to nutrient levels [GO:0031667]; response to peptide hormone [GO:0043434]; response to pyrethroid [GO:0046684]; response to salt stress [GO:0009651]; response to water deprivation [GO:0009414]; response to zinc ion [GO:0010043]; sensory perception of sound [GO:0007605]; social behavior [GO:0035176]; sphingolipid metabolic process [GO:0006665]; synaptic transmission, dopaminergic [GO:0001963]; terpene metabolic process [GO:0042214]; visual perception [GO:0007601]
138	A2RUC4	TYW5_HUMAN	TYW5 C2orf60	tRNA wybutosine-synthesizing protein 5 (hTYW5) (EC 1.14.11.42) (tRNA(Phe) (7-(3-amino-3-carboxypropyl)wyosine(37)-C(2))-hydroxylase)	H160-D162-H235	1 Fe cation	Catalytic	1.14.11.42	Unknown	No		tRNA modification [GO:0006400]; wybutosine biosynthetic process [GO:0031591]
139	O14607	UTY_HUMAN	UTY KDM6C	Histone demethylase UTY (EC 1.14.11.-) (Ubiquitously-transcribed TPR protein on the Y chromosome) (Ubiquitously-transcribed Y chromosome tetratricopeptide repeat protein)	H1093-E1095-H1173	1 Fe cation	Catalytic	1.14.11.-	Nucleus	No		regulation of gene expression [GO:0010468]

Table S5: *Functional properties of the human heme-binding proteins.*

	Uniprot Id	Entry name	Gene names	Protein names	Predicted Pattern	Types of heme cofactors	Heme role	EC number	Subcellular location	Membrane associated	Involvement in disease	Gene ontology (biological process)
1	A0A024RAI7	A0A024RAI7_HUMAN	FLJ16008 hCG_42613	FLJ16008 protein, isoform CRA_a	Unknown	heme b	Catalytic		Unknown	No		
2	Q9NP58	ABCB6_HUMAN	ABCB6 MTABC3 PRP UMAT	ATP-binding cassette sub-family B member 6, mitochondrial (Mitochondrial ABC transporter 3) (Mt-ABC transporter 3) (P-glycoprotein-related protein) (Ubiquitously-expressed mammalian ABC half transporter)	Unknown	heme b	Substrate - transport		Endoplasmic reticulum, Golgi apparatus, Mitochondrion, Cell membrane, Endosome	Yes	DISEASE: Microphthalmia, isolated, with coloboma, 7 (MCOPCB7) [MIM:614497]: A disorder of eye formation, ranging from small size of a single eye to complete bilateral absence of ocular tissues. Ocular abnormalities like opacities of the cornea and lens, scaring of the retina and choroid, and other abnormalities may also be present. Ocular colobomas are a set of malformations resulting from abnormal morphogenesis of the optic cup and stalk, and the fusion of the fetal fissure (optic fissure). {ECO:0000269|PubMed:22226084}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Dyschromatosis universalis hereditaria 3 (DUH3) [MIM:615402]: An autosomal dominant pigmentary genodermatosis characterized by a mixture of hyperpigmented and hypopigmented macules distributed randomly over the body, that appear in infancy or early childhood. The trunk and extremities are the dominant sites of abnormal pigmentation. Facial lesions can be seen in 50% of affected individuals, but involvement of palms and soles is unusual. Abnormalities of hair and nails have also been reported. Dyschromatosis universalis hereditaria may be associated with abnormalities of dermal connective tissue, nerve tissue, or other systemic complications. DISEASE: Pseudohyperkalemia, familial, 2, due to red cell leak (PSHK2) [MIM:609153]: A dominantly inherited condition characterized by increased serum potassium levels, measured in whole-blood specimens stored at or below room temperature. This condition is not accompanied by clinical symptoms or biological signs except for borderline abnormalities of red cell shape.	brain development [GO:0007420]; cellular iron ion homeostasis [GO:0006879]; heme transport [GO:0015886]; porphyrin-containing compound biosynthetic process [GO:0006779]; skin development [GO:0043588]; transmembrane transport [GO:0055085]; transport [GO:0006810]
3	O75027	ABCB7_HUMAN	ABCB7 ABC7	ATP-binding cassette sub-family B member 7, mitochondrial (ATP-binding cassette transporter 7) (ABC transporter 7 protein)	Unknown	heme b	Substrate - transport		Mitochondrion	Yes	DISEASE: Anemia, sideroblastic, spinocerebellar ataxia (ASAT) [MIM:301310]: A X-linked recessive disorder characterized by an infantile to early childhood onset of non-progressive cerebellar ataxia and mild anemia, with hypochromia and microcytosis. Note=The disease is caused by mutations affecting the gene represented in this entry.	cellular iron ion homeostasis [GO:0006879]; transmembrane transport [GO:0055085]; transport [GO:0006810]
4	Q9UNQ0	ABCG2_HUMAN	ABCG2 ABCP BCRP BCRP1 MXR	ATP-binding cassette sub-family G member 2 (Breast cancer resistance protein) (CDw338) (Mitoxantrone resistance-associated protein) (Placenta-specific ATP-binding cassette transporter) (Urate exporter) (CD antigen CD338)	Unknown	heme b	Substrate - transport		Mitochondrion, Cell membrane	Yes		cellular iron ion homeostasis [GO:0006879]; cholesterol efflux [GO:0033344]; response to drug [GO:0042493]; transport [GO:0006810]; urate metabolic process [GO:0046415]
5	Q8N7X0	ADGB_HUMAN	ADGB C6orf103 CAPN7L	Androglobin (Calpain-7-like protein)	Unknown	heme b	Oxygen storage/transport		Unknown	No		proteolysis [GO:0006508]
6	P43652	AFAM_HUMAN	AFM ALB2 ALBA	Afamin (Alpha-albumin) (Alpha-Alb)	Y377	heme b	Substrate - transport		Extracellular space	No		vitamin transport [GO:0051180]
7	P02768	ALBU_HUMAN	ALB GIG20 GIG42 PRO0903 PRO1708 PRO2044 PRO2619 PRO2675 UNQ696/PRO1341	Serum albumin	Y185	heme b	Substrate - transport		Extracellular space	No	DISEASE: Hyperthyroxinemia, familial dysalbuminemic (FDAH) [MIM:615999]: A disorder characterized by abnormally elevated levels of total serum thyroxine (T4) in euthyroid patients. It is due to abnormal serum albumin that binds T4 with enhanced affinity. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Analbuminemia (ANALBA) [MIM:616000]: A rare autosomal recessive disorder manifested by the presence of a very low amount of circulating serum albumin. Affected individuals manifest mild edema, hypotension, fatigue, and, occasionally, lower body lipodystrophy (mainly in adult females). The most common biochemical finding is hyperlipidemia, with a significant increase in the total and LDL cholesterol concentrations, but normal concentrations of HDL cholesterol and triglycerides. {ECO:0000269|PubMed:8134387}. Note=The disease is caused by mutations affecting the gene represented in this entry.	bile acid and bile salt transport [GO:0015721]; cellular protein metabolic process [GO:0044267]; cellular response to starvation [GO:0009267]; hemolysis by symbiont of host erythrocytes [GO:0019836]; high-density lipoprotein particle remodeling [GO:0034375]; maintenance of mitochondrion location [GO:0051659]; negative regulation of apoptotic process [GO:0043066]; negative regulation of programmed cell death [GO:0043069]; platelet degranulation [GO:0002576]; post-translational protein modification [GO:0043687]; receptor-mediated endocytosis [GO:0006898]; retina homeostasis [GO:0001895]; sodium-independent organic anion transport [GO:0043252]; transport [GO:0006810]

8	P02760	AMBP_HUMAN	AMBP HCP ITIL	Protein AMBP [Cleaved into: Alpha-1-microglobulin (Protein HC) (Alpha-1 microglycoprotein) (Complex-forming glycoprotein heterogeneous in charge); Inter-alpha-trypsin inhibitor light chain (ITI-LC) (Bikunin) (EDC1) (HI-30) (Uronic-acid-rich protein); Trypstatin]	Unknown	heme b	Substrate - degradation		Extracellular space	No		cell adhesion [GO:0007155]; female pregnancy [GO:0007565]; heme catabolic process [GO:0042167]; negative regulation of immune response [GO:0050777]; negative regulation of JNK cascade [GO:0046329]; protein catabolic process [GO:0030163]; protein-chromophore linkage [GO:0018298]; receptor-mediated endocytosis [GO:0006898]; viral process [GO:0016032]
9	O14867	BACH1_HUMAN	BACH1	Transcription regulator protein BACH1 (BTB and CNC homolog 1) (HA2303)	Unknown	heme b	Substrate - sensor		Nucleus	No		DNA repair [GO:0006281]; negative regulation of transcription from RNA polymerase II promoter [GO:0000122]; protein ubiquitination [GO:0016567]; regulation of transcription, DNA-templated [GO:0006355]; regulation of transcription from RNA polymerase II promoter in response to hypoxia [GO:0061418]; regulation of transcription involved in G1/S transition of mitotic cell cycle [GO:0000083]; regulation of transcription involved in G2/M transition of mitotic cell cycle [GO:0000117]
10	P15538	C11B1_HUMAN	CYP11B1 S11BH	Cytochrome P450 11B1, mitochondrial (CYPXIB1) (Cytochrome P-450c11) (Cytochrome P450C11) (Steroid 11-beta-hydroxylase) (EC 1.14.15.4)	C450	heme b	Catalytic	1.14.15.4	Mitochondrion	Yes	DISEASE: Adrenal hyperplasia 4 (AH4) [MIM:202010]: A form of congenital adrenal hyperplasia, a common recessive disease due to defective synthesis of cortisol. Congenital adrenal hyperplasia is characterized by androgen excess leading to ambiguous genitalia in affected females, rapid somatic growth during childhood in both sexes with premature closure of the epiphyses and short adult stature. Four clinical types: 'salt wasting' (SW, the most severe type), 'simple virilizing' (SV, less severely affected patients), with normal aldosterone biosynthesis, 'non-classic form' or late-onset (NC or LOAH) and 'cryptic' (asymptomatic). Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Hyperaldosteronism, familial, 1 (HALD1) [MIM:103900]: A disorder characterized by hypertension, variable hyperaldosteronism, and abnormal adrenal steroid production, including 18-oxocortisol and 18-hydroxycortisol. There is significant phenotypic heterogeneity, and some individuals never develop hypertension. Note=The disease is caused by mutations affecting the gene represented in this entry. The molecular defect causing hyperaldosteronism familial 1 is an anti-Lepore-type fusion of the CYP11B1 and CYP11B2 genes. The hybrid gene has the promoting part of CYP11B1, ACTH-sensitive, and the coding part of CYP11B2.	aldosterone biosynthetic process [GO:0032342]; C21-steroid hormone biosynthetic process [GO:0006700]; cellular response to hormone stimulus [GO:0032870]; cellular response to potassium ion [GO:0035865]; cortisol biosynthetic process [GO:0034651]; glucocorticoid biosynthetic process [GO:0006704]; glucose homeostasis [GO:0042593]; immune response [GO:0006955]; regulation of blood pressure [GO:0008217]; sterol metabolic process [GO:0016125]

11	P19099	C11B2_HUMAN	CYP11B2	Cytochrome P450 11B2, mitochondrial (Aldosterone synthase) (ALDOS) (EC 1.14.15.4) (EC 1.14.15.5) (Aldosterone-synthesizing enzyme) (CYPXIB2) (Cytochrome P-450Aldo) (Cytochrome P-450C18) (Steroid 18-hydroxylase)	C450	heme b	Catalytic	1.14.15.4; 1.14.15.5	Mitochondrion	Yes	DISEASE: Corticosterone methyloxidase 1 deficiency (CMO-1 deficiency) [MIM:203400]: Autosomal recessive disorder of aldosterone biosynthesis. There are two biochemically different forms of selective aldosterone deficiency be termed corticosterone methyloxidase (CMO) deficiency type 1 and type 2. In CMO-1 deficiency, aldosterone is undetectable in plasma, while its immediate precursor, 18-hydroxycorticosterone, is low or normal. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Corticosterone methyloxidase 2 deficiency (CMO-2 deficiency) [MIM:610600]: Autosomal recessive disorder of aldosterone biosynthesis. In CMO-2 deficiency, aldosterone can be low or normal, but at the expense of increased secretion of 18-hydroxycorticosterone. Consequently, patients have a greatly increased ratio of 18-hydroxycorticosterone to aldosterone and a low ratio of corticosterone to 18-hydroxycorticosterone in serum. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Hyperaldosteronism, familial, 1 (HALD1) [MIM:103900]: A disorder characterized by hypertension, variable hyperaldosteronism, and abnormal adrenal steroid production, including 18-oxocortisol and 18-hydroxycortisol. There is significant phenotypic heterogeneity, and some individuals never develop hypertension. Note=The disease is caused by mutations affecting the gene represented in this entry. The molecular defect causing hyperaldosteronism familial 1 is an anti-Lepore-type fusion of the CYP11B1 and CYP11B2 genes. The hybrid gene has the promoting part of CYP11B1, ACTH-sensitive, and the coding part of CYP11B2.	aldosterone biosynthetic process [GO:0032342]; C21-steroid hormone biosynthetic process [GO:0006700]; cellular response to hormone stimulus [GO:0032870]; cellular response to potassium ion [GO:0035865]; cortisol biosynthetic process [GO:0034651]; mineralocorticoid biosynthetic process [GO:0006705]; potassium ion homeostasis [GO:0055075]; regulation of blood volume by renal aldosterone [GO:0002017]; renal water homeostasis [GO:0003091]; sodium ion homeostasis [GO:0055078]; sterol metabolic process [GO:0016125]
12	Q86VB7	C163A_HUMAN	CD163 M130	Scavenger receptor cysteine-rich type 1 protein M130 (Hemoglobin scavenger receptor) (CD antigen CD163) [Cleaved into: Soluble CD163 (sCD163)]	Unknown	heme b	Substrate - degradation		Extracellular space	No		acute-phase response [GO:0006953]; receptor-mediated endocytosis [GO:0006898]
13	Q4G0S4	C27C1_HUMAN	CYP27C1	Cytochrome P450 27C1 (EC 1.14.19.-) (All-trans retinol 3,4-desaturase)	C318	heme b	Catalytic	1.14.19.-	Unknown	Yes		retinal metabolic process [GO:0042574]; retinoic acid metabolic process [GO:0042573]; retinol metabolic process [GO:0042572]
14	Q6ZSU1	C2G1P_HUMAN	CYP2G1P CYP2GP1	Putative inactive cytochrome P450 2G1 (Cytochrome P450 2G1 pseudogene)	C91	heme b	Catalytic		Unknown	No		epoxygenase P450 pathway [GO:0019373]
15	Q99643	C560_HUMAN	SDHC CYB560 SDH3	Succinate dehydrogenase cytochrome b560 subunit, mitochondrial (Integral membrane protein CII-3) (QPs-1) (QPs1) (Succinate dehydrogenase complex subunit C) (Succinate-ubiquinone oxidoreductase cytochrome B large subunit) (CYBL)	H127	heme b	Electron transfer		Mitochondrion	Yes	DISEASE: Paragangliomas 3 (PGL3) [MIM:605373]: A neural crest tumor usually derived from the chromoreceptor tissue of a paraganglion. Paragangliomas can develop at various body sites, including the head, neck, thorax and abdomen. Most commonly, they are located in the head and neck region, specifically at the carotid bifurcation, the jugular foramen, the vagal nerve, and in the middle ear. {ECO:0000269|PubMed:11062460}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Paraganglioma and gastric stromal sarcoma (PGGSS) [MIM:606864]: Gastrointestinal stromal tumors may be sporadic or inherited in an autosomal dominant manner, alone or as a component of a syndrome associated with other tumors, such as in the context of neurofibromatosis type 1 (NF1). Patients have both gastrointestinal stromal tumors and paragangliomas. Susceptibility to the tumors was inherited in an apparently autosomal dominant manner, with incomplete penetrance. {ECO:0000269|PubMed:17804857}. Note=The disease is caused by mutations affecting the gene represented in this entry.	aerobic respiration [GO:0009060]; mitochondrial electron transport, succinate to ubiquinone [GO:0006121]; oxidation-reduction process [GO:0055114]; tricarboxylic acid cycle [GO:0006099]
16	Q8N8Q1	C56D1_HUMAN	CYB561D1	Cytochrome b561 domain-containing protein 1	H55-H127; H93-H166	heme b	Electron transfer		Unknown	Yes		oxidation-reduction process [GO:0055114]
17	O14569	C56D2_HUMAN	CYB561D2 101F6 LUCA12.2	Cytochrome b561 domain-containing protein 2 (Putative tumor suppressor protein 101F6)	H48-H120;H86-H159	heme b	Electron transfer		Unknown	Yes		oxidation-reduction process [GO:0055114]

18	P04040	CATA_HUMAN	CAT	Catalase (EC 1.11.1.6)	Y358	heme b	Catalytic	1.11.1.6	Peroxisome	No	DISEASE: Acatalasemia (ACATLAS) [MIM:614097]: A metabolic disorder characterized by a total or near total loss of catalase activity in red cells. It is often associated with ulcerating oral lesions. {ECO:0000269|PubMed:2308162}. Note=The disease is caused by mutations affecting the gene represented in this entry.	aerobic respiration [GO:0009060]; aging [GO:0007568]; cellular response to growth factor stimulus [GO:0071363]; cellular response to oxidative stress [GO:0034599]; cholesterol metabolic process [GO:0008203]; hemoglobin metabolic process [GO:0020027]; hydrogen peroxide catabolic process [GO:0042744]; negative regulation of apoptotic process [GO:0043066]; response to cadmium ion [GO:0046686]; response to drug [GO:0042493]; response to estradiol [GO:0032355]; response to ethanol [GO:0045471]; response to fatty acid [GO:0070542]; response to hydrogen peroxide [GO:0042542]; response to hyperoxia [GO:0055093]; response to hypoxia [GO:0001666]; response to inactivity [GO:0014854]; response to insulin [GO:0032868]; response to L-ascorbic acid [GO:0033591]; response to lead ion [GO:0010288]; response to light intensity [GO:0009642]; response to ozone [GO:0010193]; response to phenylpropanoid [GO:0080184]; response to reactive oxygen species [GO:0000302]; response to vitamin A [GO:0033189]; response to vitamin E [GO:0033197]; triglyceride metabolic process [GO:0006641]; ureteric bud development [GO:0001657]; UV protection [GO:0009650]
19	Q6P9G0	CB5D1_HUMAN	CYB5D1	Cytochrome b5 domain-containing protein 1	Y52-H83	heme b	Electron transfer		Unknown	No		
20	P35520	CBS_HUMAN	CBS	Cystathionine beta-synthase (EC 4.2.1.22) (Beta-thionase) (Serine sulfhydrase)	C52-H65	heme b	Regulatory - catalysis	4.2.1.22	Cytoplasm, Nucleus	No	DISEASE: Cystathionine beta-synthase deficiency (CBSD) [MIM:236200]: An enzymatic deficiency resulting in altered sulfur metabolism and homocystinuria. The clinical features of untreated homocystinuria due to CBS deficiency include myopia, ectopia lentis, mental retardation, skeletal anomalies resembling Marfan syndrome, and thromboembolic events. Light skin and hair can also be present. Biochemical features include increased urinary homocystine and methionine. Note=The disease is caused by mutations affecting the gene represented in this entry.	cysteine biosynthetic process [GO:0019344]; cysteine biosynthetic process from serine [GO:0006535]; cysteine biosynthetic process via cystathionine [GO:0019343]; DNA protection [GO:0042262]; homocysteine catabolic process [GO:0043418]; homocysteine metabolic process [GO:0050667]; hydrogen sulfide biosynthetic process [GO:0070814]; L-cysteine catabolic process [GO:0019448]; L-serine catabolic process [GO:0006565]; L-serine metabolic process [GO:0006563]; transsulfuration [GO:0019346]
21	P0DN79	CBSL_HUMAN	CBSL	Cystathionine beta-synthase-like protein (EC 4.2.1.22) (Beta-thionase) (Serine sulfhydrase)	C52-H65	heme b	Unknown	4.2.1.22	Cytoplasm, Nucleus	No		cysteine biosynthetic process from serine [GO:0006535]; cysteine biosynthetic process via cystathionine [GO:0019343]
22	P53701	CCHL_HUMAN	HCCS CCHL	Cytochrome c-type heme lyase (CCHL) (EC 4.4.1.17) (Holocytochrome c-type synthase)	Unknown	heme c	Substrate - Protein biosynthesis	4.4.1.17	Mitochondrion	Yes	DISEASE: Linear skin defects with multiple congenital anomalies 1 (LSDMCA1) [MIM:309801]: A disorder characterized by dermal, ocular, neurological and cardiac abnormalities. LSDMCA1 main features are unilateral or bilateral microphthalmia, linear skin defects in affected females, and in utero lethality for males. Skin defects are limited to the face and neck, consisting of areas of aplastic skin that heal with age to form hyperpigmented areas. Additional features in female patients include agenesis of the corpus callosum, sclerocornea, chorioretinal abnormalities, infantile seizures, congenital heart defect, mental retardation, and diaphragmatic hernia. Microphthalmia is a disorder of eye formation, ranging from small size of a single eye to complete bilateral absence of ocular tissues (anophthalmia). In many cases, microphthalmia/anophthalmia occurs in association with syndromes that include non-ocular abnormalities. {ECO:0000269|PubMed:17033964}. Note=The disease is caused by mutations affecting the gene represented in this entry.	animal organ morphogenesis [GO:0009887]; cytochrome c-heme linkage [GO:0018063]; oxidation-reduction process [GO:0055114]
23	P53621	COPA_HUMAN	COPA	Coatomer subunit alpha (Alpha-coat protein) (Alpha-COP) (HEP-COP) (HEPCOP) [Cleaved into: Xenin (Xenopsin-related peptide); Proxenin]	Unknown	heme d1	Catalytic		Cytoplasm, Golgi apparatus	Yes	DISEASE: Autoimmune interstitial lung, joint, and kidney disease (AILJK) [MIM:616414]: An autoimmune disease characterized by inflammatory arthritis, interstitial lung disease, and immune complex-mediated renal disease. {ECO:0000269|PubMed:25894502}. Note=The disease is caused by mutations affecting the gene represented in this entry.	ER to Golgi vesicle-mediated transport [GO:0006888]; intracellular protein transport [GO:0006886]; intra-Golgi vesicle-mediated transport [GO:0006891]; pancreatic juice secretion [GO:0030157]; retrograde vesicle-mediated transport, Golgi to ER [GO:0006890]

24	P00395	COX1_HUMAN	MT-CO1 COI COXI MTCO1	Cytochrome c oxidase subunit 1 (EC 1.9.3.1) (Cytochrome c oxidase polypeptide I)	H61-H378; H328-H376	heme a, heme a3	Electron transfer	1.9.3.1	Mitochondrion	Yes	DISEASE: Leber hereditary optic neuropathy (LHON) [MIM:535000]: A maternally inherited disease resulting in acute or subacute loss of central vision, due to optic nerve dysfunction. Cardiac conduction defects and neurological defects have also been described in some patients. LHON results from primary mitochondrial DNA mutations affecting the respiratory chain complexes. {ECO:0000269|PubMed:1322638}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Note=MT-CO1 may play a role in the pathogenesis of acquired idiopathic sideroblastic anemia, a disease characterized by inadequate formation of heme and excessive accumulation of iron in mitochondria. Mitochondrial iron overload may be attributable to mutations of mitochondrial DNA because these can cause respiratory chain dysfunction, thereby impairing reduction of ferric iron to ferrous iron. The reduced form of iron is essential to the last step of mitochondrial heme biosynthesis. {ECO:0000269|PubMed:9389715, ECO:0000269|PubMed:9851701}.; DISEASE: Mitochondrial complex IV deficiency (MT-C4D) [MIM:220110]: A disorder of the mitochondrial respiratory chain with heterogeneous clinical manifestations, ranging from isolated myopathy to severe multisystem disease affecting several tissues and organs. Features include hypertrophic cardiomyopathy, hepatomegaly and liver dysfunction, hypotonia, muscle weakness, exercise intolerance, developmental delay, delayed motor development and mental retardation. Some affected individuals manifest a fatal hypertrophic cardiomyopathy resulting in neonatal death. A subset of patients manifest Leigh syndrome. {ECO:0000269|PubMed:12140182, ECO:0000269|PubMed:16284789}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Recurrent myoglobinuria mitochondrial (RM-MT) [MIM:550500]: Recurrent myoglobinuria is characterized by recurrent attacks of rhabdomyolysis (necrosis or disintegration of skeletal muscle) associated with muscle pain and weakness, and followed by excretion of myoglobin in the urine. {ECO:0000269|PubMed:10980727}. Note=The gene represented in this entry may be involved in disease pathogenesis.; DISEASE: Deafness, sensorineural, mitochondrial (DFNM) [MIM:500008]: A form of non-syndromic deafness with maternal inheritance. Affected individuals manifest progressive, postlingual, sensorineural hearing loss involving high frequencies. {ECO:0000269|PubMed:10577941}. Note=The disease is caused by mutations affecting the gene represented in this entry.;	aerobic respiration [GO:0009060]; aging [GO:0007568]; cerebellum development [GO:0021549]; electron transport coupled proton transport [GO:0015990]; mitochondrial electron transport, cytochrome c to oxygen [GO:0006123]; response to copper ion [GO:0046688]; response to electrical stimulus [GO:0051602]; response to oxidative stress [GO:0006979]

25	Q7KZN9	COX15_HUMAN	COX15	Cytochrome c oxidase assembly protein COX15 homolog	Unknown	heme o	Substrate - modification		Mitochondrion	Yes	DISEASE: Cardioencephalomyopathy, fatal infantile, due to cytochrome c oxidase deficiency 2 (CEMCOX2) [MIM:615119]: An infantile disorder, with a rapidly progressive fatal course, characterized by cytochrome c oxidase deficiency. Clinical features include microcephaly, encephalopathy, hypertrophic cardiomyopathy, persistent lactic acidosis, respiratory distress, hypotonia and seizures. Postmortem cardiac muscle studies show marked complex IV deficiency. Complex IV activity is only slightly decreased in the skeletal muscle. {ECO:0000269|PubMed:12474143, ECO:0000269|PubMed:21412973}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Leigh syndrome (LS) [MIM:256000]: An early-onset progressive neurodegenerative disorder characterized by the presence of focal, bilateral lesions in one or more areas of the central nervous system including the brainstem, thalamus, basal ganglia, cerebellum and spinal cord. Clinical features depend on which areas of the central nervous system are involved and include subacute onset of psychomotor retardation, hypotonia, ataxia, weakness, vision loss, eye movement abnormalities, seizures, and dysphagia. {ECO:0000269|PubMed:15235026, ECO:0000269|PubMed:15863660}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cellular respiration [GO:0045333]; heme a biosynthetic process [GO:0006784]; heme biosynthetic process [GO:0006783]; hydrogen ion transmembrane transport [GO:1902600]; mitochondrial electron transport, cytochrome c to oxygen [GO:0006123]; oxidation-reduction process [GO:0055114]; respiratory chain complex IV assembly [GO:0008535]; respiratory gaseous exchange [GO:0007585]
26	P20674	COX5A_HUMAN	COX5A	Cytochrome c oxidase subunit 5A, mitochondrial (Cytochrome c oxidase polypeptide Va)	Unknown	heme a	Catalytic		Mitochondrion	Yes	DISEASE: Note=Mitochondrial complex IV deficiency is a rare condition caused by mutation in COX5A that lead to pulmonary arterial hypertension (PAH), failure to thrive and lactic acidemia. {ECO:0000269|PubMed:28247525}.	mitochondrial electron transport, cytochrome c to oxygen [GO:0006123]
27	Q8NHV5	CP052_HUMAN	C16orf52	Uncharacterized protein C16orf52	E125	heme b	Unknown		Unknown	No		
28	P05108	CP11A_HUMAN	CYP11A1 CYP11A	Cholesterol side-chain cleavage enzyme, mitochondrial (EC 1.14.15.6) (CYPXIA1) (Cholesterol desmolase) (Cytochrome P450 11A1) (Cytochrome P450(scc))	C462	heme b	Catalytic	1.14.15.6	Mitochondrion	Yes	DISEASE: Adrenal insufficiency, congenital, with 46,XY sex reversal (AICSR) [MIM:613743]: A rare disorder that can present as acute adrenal insufficiency in infancy or childhood. ACTH and plasma renin activity are elevated and adrenal steroids are inappropriately low or absent; the 46,XY patients have female external genitalia, sometimes with clitoromegaly. The phenotypic spectrum ranges from prematurity, complete underandrogenization, and severe early-onset adrenal failure to term birth with clitoromegaly and later-onset adrenal failure. Patients with congenital adrenal insufficiency do not manifest the massive adrenal enlargement typical of congenital lipoid adrenal hyperplasia. Note=The disease is caused by mutations affecting the gene represented in this entry.	C21-steroid hormone biosynthetic process [GO:0006700]; cholesterol metabolic process [GO:0008203]; sterol metabolic process [GO:0016125]; vitamin D metabolic process [GO:0042359]
29	P05093	CP17A_HUMAN	CYP17A1 CYP17 S17AH	Steroid 17-alpha-hydroxylase/17,20 lyase (EC 1.14.14.19) (17-alpha-hydroxyprogesterone aldolase) (EC 1.14.14.32) (CYPXVII) (Cytochrome P450 17A1) (Cytochrome P450-C17) (Cytochrome P450c17) (Steroid 17-alpha-monooxygenase)	C442	heme b	Catalytic	1.14.14.19; 1.14.14.32	Unknown	Yes	DISEASE: Adrenal hyperplasia 5 (AH5) [MIM:202110]: A form of congenital adrenal hyperplasia, a common recessive disease due to defective synthesis of cortisol. Congenital adrenal hyperplasia is characterized by androgen excess leading to ambiguous genitalia in affected females, rapid somatic growth during childhood in both sexes with premature closure of the epiphyses and short adult stature. Four clinical types: 'salt wasting' (SW, the most severe type), 'simple virilizing' (SV, less severely affected patients), with normal aldosterone biosynthesis, 'non-classic form' or late-onset (NC or LOAH) and 'cryptic' (asymptomatic). Note=The disease is caused by mutations affecting the gene represented in this entry.	androgen biosynthetic process [GO:0006702]; glucocorticoid biosynthetic process [GO:0006704]; hormone biosynthetic process [GO:0042446]; progesterone metabolic process [GO:0042448]; sex differentiation [GO:0007548]; steroid biosynthetic process [GO:0006694]; steroid metabolic process [GO:0008202]; sterol metabolic process [GO:0016125]

30	P11511	CP19A_HUMAN	CYP19A1 ARO1 CYAR CYP19	Aromatase (EC 1.14.14.14) (CYPXIX) (Cytochrome P-450AROM) (Cytochrome P450 19A1) (Estrogen synthase)	C437	heme b	Catalytic	1.14.14.14	Unknown	Yes	DISEASE: Aromatase excess syndrome (AEXS) [MIM:139300]: An autosomal dominant disorder characterized by increased extraglandular aromatization of steroids that presents with heterosexual precocity in males and isosexual precocity in females. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Aromatase deficiency (AROD) [MIM:613546]: A rare disease in which fetal androgens are not converted into estrogens due to placental aromatase deficiency. Thus, pregnant women exhibit a hirsutism, which spontaneously resolves after post-partum. At birth, female babies present with pseudohermaphroditism due to virilization of extern genital organs. In adult females, manifestations include delay of puberty, breast hypoplasia and primary amenorrhoea with multicystic ovaries. Note=The disease is caused by mutations affecting the gene represented in this entry.	androgen catabolic process [GO:0006710]; estrogen biosynthetic process [GO:0006703]; female genitalia development [GO:0030540]; female gonad development [GO:0008585]; mammary gland development [GO:0030879]; negative regulation of chronic inflammatory response [GO:0002677]; negative regulation of macrophage chemotaxis [GO:0010760]; positive regulation of estradiol secretion [GO:2000866]; prostate gland growth [GO:0060736]; steroid biosynthetic process [GO:0006694]; sterol metabolic process [GO:0016125]; testosterone biosynthetic process [GO:0061370]; uterus development [GO:0060065]
31	P04798	CP1A1_HUMAN	CYP1A1	Cytochrome P450 1A1 (EC 1.14.14.1) (CYPIA1) (Cytochrome P450 form 6) (Cytochrome P450-C) (Cytochrome P450-P1)	C457	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes		cellular response to copper ion [GO:0071280]; cellular response to organic cyclic compound [GO:0071407]; coumarin metabolic process [GO:0009804]; dibenzo-p-dioxin catabolic process [GO:0019341]; digestive tract development [GO:0048565]; drug metabolic process [GO:0017144]; epoxygenase P450 pathway [GO:0019373]; ethylene metabolic process [GO:0009692]; flavonoid metabolic process [GO:0009812]; hepatocyte differentiation [GO:0070365]; hydrogen peroxide biosynthetic process [GO:0050665]; insecticide metabolic process [GO:0017143]; lipid hydroxylation [GO:0002933]; maternal process involved in parturition [GO:0060137]; omega-hydroxylase P450 pathway [GO:0097267]; regulation of lipid metabolic process [GO:0019216]; response to antibiotic [GO:0046677]; response to arsenic-containing substance [GO:0046685]; response to drug [GO:0042493]; response to food [GO:0032094]; response to herbicide [GO:0009635]; response to hyperoxia [GO:0055093]; response to hypoxia [GO:0001666]; response to immobilization stress [GO:0035902]; response to iron(III) ion [GO:0010041]; response to lipopolysaccharide [GO:0032496]; response to nematode [GO:0009624]; response to virus [GO:0009615]; response to vitamin A [GO:0033189]; response to wounding [GO:0009611]; steroid metabolic process [GO:0008202]; vitamin D metabolic process [GO:0042359]
32	P05177	CP1A2_HUMAN	CYP1A2	Cytochrome P450 1A2 (EC 1.14.14.1) (CYPIA2) (Cholesterol 25-hydroxylase) (Cytochrome P(3)450) (Cytochrome P450 4) (Cytochrome P450-P3)	C458	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes		alkaloid metabolic process [GO:0009820]; cellular respiration [GO:0045333]; cellular response to cadmium ion [GO:0071276]; cellular response to copper ion [GO:0071280]; dibenzo-p-dioxin metabolic process [GO:0018894]; drug catabolic process [GO:0042737]; drug metabolic process [GO:0017144]; epoxygenase P450 pathway [GO:0019373]; exogenous drug catabolic process [GO:0042738]; heterocycle metabolic process [GO:0046483]; hydrogen peroxide biosynthetic process [GO:0050665]; lung development [GO:0030324]; methylation [GO:0032259]; monocarboxylic acid metabolic process [GO:0032787]; monoterpenoid metabolic process [GO:0016098]; omega-hydroxylase P450 pathway [GO:0097267]; oxidation-reduction process [GO:0055114]; oxidative deethylation [GO:0071615]; oxidative demethylation [GO:0070989]; porphyrin-containing compound metabolic process [GO:0006778]; post-embryonic development [GO:0009791]; regulation of gene expression [GO:0010468]; response to estradiol [GO:0032355]; response to immobilization stress [GO:0035902]; response to lipopolysaccharide [GO:0032496]; steroid catabolic process [GO:0006706]; toxin biosynthetic process [GO:0009403]; xenobiotic metabolic process [GO:0006805]

33	Q16678	CP1B1_HUMAN	CYP1B1	Cytochrome P450 1B1 (EC 1.14.14.1) (CYPIB1)	C470	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum, Mitochondrion	Yes	DISEASE: Anterior segment dysgenesis 6 (ASGD6) [MIM:617315]: A form of anterior segment dysgenesis, a group of defects affecting anterior structures of the eye including cornea, iris, lens, trabecular meshwork, and Schlemm canal; DISEASE: Glaucoma 3, primary congenital, A (GLC3A) [MIM:231300]: An autosomal recessive form of primary congenital glaucoma (PCG). DISEASE: Glaucoma, primary open angle (POAG) [MIM:137760]: A complex and genetically heterogeneous ocular disorder characterized by a specific pattern of optic nerve and visual field defects. Note=Disease susceptibility is associated with variations affecting the gene represented in this entry. CYP1B1 mutations have been reported to pose a significant risk for early-onset POAG and also modify glaucoma phenotype in patients who do not carry a MYOC mutation (PubMed:15342693). {ECO:0000269|PubMed:15342693}.; DISEASE: Glaucoma 1, open angle, A (GLC1A) [MIM:137750]: A form of primary open angle glaucoma (POAG). POAG is characterized by a specific pattern of optic nerve and visual field defects. The angle of the anterior chamber of the eye is open, and usually the intraocular pressure is increased. However, glaucoma can occur at any intraocular pressure. The disease is generally asymptomatic until the late stages, by which time significant and irreversible optic nerve damage has already taken place. {ECO:0000269|PubMed:11774072}. Note=The gene represented in this entry acts as a disease modifier. Digenic mutations in CYP1B1 and MYOC have been found in a family segregating both primary adult-onset and juvenile forms of open angle glaucoma (PubMed:11774072). All affected family members with mutations in both MYOC and CYP1B1 had juvenile glaucoma, whereas those with only the MYOC mutation had the adult-onset form (PubMed:11774072).	angiogenesis [GO:0001525]; cellular response to hydrogen peroxide [GO:0070301]; cellular response to organic cyclic compound [GO:0071407]; collagen fibril organization [GO:0030199]; endothelial cell-cell adhesion [GO:0071603]; endothelial cell migration [GO:0043542]; epoxygenase P450 pathway [GO:0019373]; estrogen metabolic process [GO:0008210]; intrinsic apoptotic signaling pathway in response to oxidative stress [GO:0008631]; membrane lipid catabolic process [GO:0046466]; negative regulation of cell adhesion mediated by integrin [GO:0033629]; negative regulation of cell migration [GO:0030336]; negative regulation of cell proliferation [GO:0008285]; negative regulation of NF-kappaB transcription factor activity [GO:0032088]; nitric oxide biosynthetic process [GO:0006809]; omega-hydroxylase P450 pathway [GO:0097267]; oxidation-reduction process [GO:0055114]; positive regulation of angiogenesis [GO:0045766]; positive regulation of apoptotic process [GO:0043065]; positive regulation of JAK-STAT cascade [GO:0046427]; positive regulation of vascular endothelial growth factor production [GO:0010575]; regulation of reactive oxygen species metabolic process [GO:2000377]; response to toxic substance [GO:0009636]; retinal blood vessel morphogenesis [GO:0061304]; retinal metabolic process [GO:0042574]; sterol metabolic process [GO:0016125]; toxin metabolic process [GO:0009404]; xenobiotic metabolic process [GO:0006805]
34	Q6UW02	CP20A_HUMAN	CYP20A1 UNQ667/PRO1301	Cytochrome P450 20A1 (EC 1.14.-.-)	C409	heme b	Catalytic	1.14.-.-	Unknown	Yes		
35	P08686	CP21A_HUMAN	CYP21A2 CYP21 CYP21B	Steroid 21-hydroxylase (EC 1.14.14.16) (21-OHase) (Cytochrome P-450c21) (Cytochrome P450 21) (Cytochrome P450 XXI) (Cytochrome P450-C21) (Cytochrome P450-C21B)	C428	heme b	Catalytic	1.14.14.16	Endoplasmic reticulum	Yes	DISEASE: Adrenal hyperplasia 3 (AH3) [MIM:201910]: A form of congenital adrenal hyperplasia, a common recessive disease due to defective synthesis of cortisol. Congenital adrenal hyperplasia is characterized by androgen excess leading to ambiguous genitalia in affected females, rapid somatic growth during childhood in both sexes with premature closure of the epiphyses and short adult stature. Four clinical types: 'salt wasting' (SW, the most severe type), 'simple virilizing' (SV, less severely affected patients), with normal aldosterone biosynthesis, 'non-classic form' or late-onset (NC or LOAH) and 'cryptic' (asymptomatic). Note=The disease is caused by mutations affecting the gene represented in this entry.	glucocorticoid biosynthetic process [GO:0006704]; mineralocorticoid biosynthetic process [GO:0006705]; steroid biosynthetic process [GO:0006694]; steroid metabolic process [GO:0008202]; sterol metabolic process [GO:0016125]
36	Q07973	CP24A_HUMAN	CYP24A1 CYP24	1,25-dihydroxyvitamin D(3) 24-hydroxylase, mitochondrial (24-OHase) (Vitamin D(3) 24-hydroxylase) (EC 1.14.15.16) (Cytochrome P450 24A1) (Cytochrome P450-CC24)	C462	heme b	Catalytic	1.14.15.16	Mitochondrion	No	DISEASE: Hypercalcemia, infantile, 1 (HCINF1) [MIM:143880]: A disorder characterized by abnormally high level of calcium in the blood, failure to thrive, vomiting, dehydration, and nephrocalcinosis. {ECO:0000269|PubMed:21675912}. Note=The disease is caused by mutations affecting the gene represented in this entry.	osteoblast differentiation [GO:0001649]; oxidation-reduction process [GO:0055114]; response to vitamin D [GO:0033280]; vitamin D catabolic process [GO:0042369]; vitamin D metabolic process [GO:0042359]; vitamin D receptor signaling pathway [GO:0070561]; vitamin metabolic process [GO:0006766]
37	O43174	CP26A_HUMAN	CYP26A1 CYP26 P450RAI1	Cytochrome P450 26A1 (EC 1.14.13.-) (Cytochrome P450 retinoic acid-inactivating 1) (Cytochrome P450RAI) (hP450RAI) (Retinoic acid 4-hydroxylase) (Retinoic acid-metabolizing cytochrome)	H133-C442	heme b	Catalytic	1.14.13.-	Endoplasmic reticulum	Yes		negative regulation of retinoic acid receptor signaling pathway [GO:0048387]; retinoic acid catabolic process [GO:0034653]; retinoic acid metabolic process [GO:0042573]; sterol metabolic process [GO:0016125]; vitamin metabolic process [GO:0006766]; xenobiotic metabolic process [GO:0006805]

38	Q9NR63	CP26B_HUMAN	CYP26B1 CYP26A2 P450RAI2	Cytochrome P450 26B1 (EC 1.14.13.-) (Cytochrome P450 26A2) (Cytochrome P450 retinoic acid-inactivating 2) (Cytochrome P450RAI-2) (Retinoic acid-metabolizing cytochrome)	H138-C441	heme b	Catalytic	1.14.13.-	Endoplasmic reticulum	Yes	DISEASE: Radiohumeral fusions with other skeletal and craniofacial anomalies (RHFCA) [MIM:614416]: A disease characterized by craniofacial malformations, occipital encephalocele, radiohumeral fusions, oligodactyly, advanced osseous maturation, and calvarial mineralization defects. {ECO:0000269|PubMed:22019272}. Note=The disease is caused by mutations affecting the gene represented in this entry.	bone morphogenesis [GO:0060349]; cell fate determination [GO:0001709]; cellular response to retinoic acid [GO:0071300]; cornification [GO:0070268]; embryonic limb morphogenesis [GO:0030326]; establishment of skin barrier [GO:0061436]; establishment of T cell polarity [GO:0001768]; inflammatory response [GO:0006954]; male meiotic nuclear division [GO:0007140]; negative regulation of retinoic acid receptor signaling pathway [GO:0048387]; oxidation-reduction process [GO:0055114]; positive regulation of gene expression [GO:0010628]; positive regulation of tongue muscle cell differentiation [GO:2001037]; proximal/distal pattern formation [GO:0009954]; regulation of T cell differentiation [GO:0045580]; retinoic acid catabolic process [GO:0034653]; retinoic acid receptor signaling pathway [GO:0048384]; spermatogenesis [GO:0007283]; sterol metabolic process [GO:0016125]; tongue morphogenesis [GO:0043587]; vitamin metabolic process [GO:0006766]; xenobiotic metabolic process [GO:0006805]
39	Q6V0L0	CP26C_HUMAN	CYP26C1	Cytochrome P450 26C1 (EC 1.14.-.-)	H138-C459	heme b	Catalytic	1.14.-.-	Unknown	Yes	DISEASE: Focal facial dermal dysplasia 4 (FFDD4) [MIM:614974]: A form of focal facial dermal dysplasia, a group of developmental defects characterized by bitemporal or preauricular skin lesions resembling aplasia cutis congenita. Skin defects occur at the sites of facial fusion during embryogenesis, with temporal lesions situated at the junction between the frontonasal and maxillary facial prominences, and preauricular lesions at the meeting point of the maxillary and mandibular prominences. The ectodermal lesions show consistent histologic abnormalities: atrophy and flattening of the epidermis, replacement of the dermis by loose connective tissue, reduced levels of fragmented elastic tissue and absence of the subcutaneous tissues and adnexal structures. FFDD4 is characterized by isolated, preauricular skin lesions. {ECO:0000269|PubMed:23161670}. Note=The disease is caused by mutations affecting the gene represented in this entry.	anterior/posterior pattern specification [GO:0009952]; central nervous system development [GO:0007417]; negative regulation of retinoic acid receptor signaling pathway [GO:0048387]; neural crest cell development [GO:0014032]; organelle fusion [GO:0048284]; oxidation-reduction process [GO:0055114]; retinoic acid catabolic process [GO:0034653]; sterol metabolic process [GO:0016125]; vitamin metabolic process [GO:0006766]
40	Q02318	CP27A_HUMAN	CYP27A1 CYP27	Sterol 26-hydroxylase, mitochondrial (EC 1.14.15.15) (5-beta-cholestane-3-alpha,7-alpha,12-alpha-triol 27-hydroxylase) (Cytochrome P-450C27/25) (Cytochrome P450 27) (Sterol 27-hydroxylase) (Vitamin D(3) 25-hydroxylase)	C476	heme b	Catalytic	1.14.15.15	Mitochondrion	Yes	DISEASE: Cerebrotendinous xanthomatosis (CTX) [MIM:213700]: Rare sterol storage disorder characterized clinically by progressive neurologic dysfunction, premature atherosclerosis, and cataracts. Note=The disease is caused by mutations affecting the gene represented in this entry.	bile acid biosynthetic process [GO:0006699]; sterol metabolic process [GO:0016125]
41	O15528	CP27B_HUMAN	CYP27B1 CYP1ALPHA CYP27B	25-hydroxyvitamin D-1 alpha hydroxylase, mitochondrial (EC 1.14.15.18) (25-OHD-1 alpha-hydroxylase) (25-hydroxyvitamin D(3) 1-alpha-hydroxylase) (VD3 1A hydroxylase) (Calcidiol 1-monooxygenase) (Cytochrome P450 subfamily XXVIIB polypeptide 1) (Cytochrome P450C1 alpha) (Cytochrome P450VD1-alpha) (Cytochrome p450 27B1)	C455	heme b	Catalytic	1.14.15.18	Mitochondrion	Yes	DISEASE: Rickets vitamin D-dependent 1A (VDDR1A) [MIM:264700]: A disorder caused by a selective deficiency of the active form of vitamin D (1,25-dihydroxyvitamin D3) and resulting in defective bone mineralization and clinical features of rickets. Note=The disease is caused by mutations affecting the gene represented in this entry.	bone mineralization [GO:0030282]; calcitriol biosynthetic process from calciol [GO:0036378]; calcium ion homeostasis [GO:0055074]; calcium ion transport [GO:0006816]; decidualization [GO:0046697]; G1 to G0 transition [GO:0070314]; negative regulation of calcidiol 1-monooxygenase activity [GO:0010956]; negative regulation of cell growth [GO:0030308]; negative regulation of cell proliferation [GO:0008285]; positive regulation of keratinocyte differentiation [GO:0045618]; positive regulation of vitamin D 24-hydroxylase activity [GO:0010980]; positive regulation of vitamin D receptor signaling pathway [GO:0070564]; regulation of bone mineralization [GO:0030500]; response to estrogen [GO:0043627]; response to interferon-gamma [GO:0034341]; response to lipopolysaccharide [GO:0032496]; response to vitamin D [GO:0033280]; vitamin D catabolic process [GO:0042369]; vitamin D metabolic process [GO:0042359]; vitamin metabolic process [GO:0006766]
42	P11509	CP2A6_HUMAN	CYP2A6 CYP2A3	Cytochrome P450 2A6 (EC 1.14.13.-) (1,4-cineole 2-exo-monooxygenase) (CYPIIA6) (Coumarin 7-hydroxylase) (Cytochrome P450 IIA3) (Cytochrome P450(I))	C439	heme b	Catalytic	1.14.13.-	Endoplasmic reticulum	Yes		coumarin catabolic process [GO:0046226]; coumarin metabolic process [GO:0009804]; drug metabolic process [GO:0017144]; epoxygenase P450 pathway [GO:0019373]; exogenous drug catabolic process [GO:0042738]; steroid metabolic process [GO:0008202]
43	P20853	CP2A7_HUMAN	CYP2A7	Cytochrome P450 2A7 (EC 1.14.14.1) (CYPIIA7) (Cytochrome P450 IIA4)	C439	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes		epoxygenase P450 pathway [GO:0019373]
44	Q16696	CP2AD_HUMAN	CYP2A13	Cytochrome P450 2A13 (EC 1.14.14.1) (CYPIIA13)	C439	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes		coumarin metabolic process [GO:0009804]; epoxygenase P450 pathway [GO:0019373]; xenobiotic metabolic process [GO:0006805]

45	P20813	CP2B6_HUMAN	CYP2B6	Cytochrome P450 2B6 (EC 1.14.13.-) (1,4-cineole 2-exo-monooxygenase) (CYPIIB6) (Cytochrome P450 IIB1)	C436	heme b	Catalytic	1.14.13.-	Endoplasmic reticulum	Yes		cellular ketone metabolic process [GO:0042180]; drug metabolic process [GO:0017144]; epoxygenase P450 pathway [GO:0019373]; exogenous drug catabolic process [GO:0042738]; oxidation-reduction process [GO:0055114]; steroid metabolic process [GO:0008202]; xenobiotic metabolic process [GO:0006805]
46	P10632	CP2C8_HUMAN	CYP2C8	Cytochrome P450 2C8 (EC 1.14.14.1) (CYPIIC8) (Cytochrome P450 IIC2) (Cytochrome P450 MP-12) (Cytochrome P450 MP-20) (Cytochrome P450 form 1) (S-mephenytoin 4-hydroxylase)	C435	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes		drug metabolic process [GO:0017144]; epoxygenase P450 pathway [GO:0019373]; exogenous drug catabolic process [GO:0042738]; lipid hydroxylation [GO:0002933]; omega-hydroxylase P450 pathway [GO:0097267]; organic acid metabolic process [GO:0006082]; oxidation-reduction process [GO:0055114]; oxidative demethylation [GO:0070989]; steroid metabolic process [GO:0008202]; xenobiotic metabolic process [GO:0006805]
47	P11712	CP2C9_HUMAN	CYP2C9 CYP2C10	Cytochrome P450 2C9 (EC 1.14.13.-) ((R)-limonene 6-monooxygenase) (EC 1.14.14.53) ((S)-limonene 6-monooxygenase) (EC 1.14.14.51) ((S)-limonene 7-monooxygenase) (EC 1.14.14.52) (CYPIIC9) (Cholesterol 25-hydroxylase) (EC 1.14.99.38) (Cytochrome P-450MP) (Cytochrome P450 MP-4) (Cytochrome P450 MP-8) (Cytochrome P450 PB-1) (S-mephenytoin 4-hydroxylase)	C435	heme b	Catalytic	1.14.13.-; 1.14.14.53; 1.14.14.51; 1.14.14.52; 1.14.99.38	Endoplasmic reticulum	Yes		cellular amide metabolic process [GO:0043603]; drug catabolic process [GO:0042737]; drug metabolic process [GO:0017144]; epoxygenase P450 pathway [GO:0019373]; exogenous drug catabolic process [GO:0042738]; monocarboxylic acid metabolic process [GO:0032787]; monoterpenoid metabolic process [GO:0016098]; omega-hydroxylase P450 pathway [GO:0097267]; oxidation-reduction process [GO:0055114]; oxidative demethylation [GO:0070989]; steroid metabolic process [GO:0008202]; urea metabolic process [GO:0019627]; xenobiotic metabolic process [GO:0006805]
48	P33260	CP2CI_HUMAN	CYP2C18	Cytochrome P450 2C18 (EC 1.14.14.1) (CYPIIC18) (Cytochrome P450-6b/29c)	C435	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes		epoxygenase P450 pathway [GO:0019373]; xenobiotic metabolic process [GO:0006805]
49	P33261	CP2CJ_HUMAN	CYP2C19	Cytochrome P450 2C19 (EC 1.14.13.-) ((R)-limonene 6-monooxygenase) (EC 1.14.14.53) ((S)-limonene 6-monooxygenase) (EC 1.14.14.51) ((S)-limonene 7-monooxygenase) (EC 1.14.14.52) (CYPIIC17) (CYPIIC19) (Cytochrome P450-11A) (Cytochrome P450-254C) (Mephenytoin 4-hydroxylase)	C435	heme b	Catalytic	1.14.13.-; 1.14.14.53; 1.14.14.51; 1.14.14.52	Endoplasmic reticulum	Yes		drug metabolic process [GO:0017144]; epoxygenase P450 pathway [GO:0019373]; exogenous drug catabolic process [GO:0042738]; heterocycle metabolic process [GO:0046483]; monoterpenoid metabolic process [GO:0016098]; omega-hydroxylase P450 pathway [GO:0097267]; oxidation-reduction process [GO:0055114]; steroid metabolic process [GO:0008202]; xenobiotic metabolic process [GO:0006805]
50	P10635	CP2D6_HUMAN	CYP2D6 CYP2DL1	Cytochrome P450 2D6 (EC 1.14.14.1) (CYPIID6) (Cholesterol 25-hydroxylase) (Cytochrome P450-DB1) (Debrisoquine 4-hydroxylase)	C443	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes		alkaloid catabolic process [GO:0009822]; alkaloid metabolic process [GO:0009820]; arachidonic acid metabolic process [GO:0019369]; coumarin metabolic process [GO:0009804]; drug catabolic process [GO:0042737]; drug metabolic process [GO:0017144]; heterocycle metabolic process [GO:0046483]; isoquinoline alkaloid metabolic process [GO:0033076]; monoterpenoid metabolic process [GO:0016098]; negative regulation of binding [GO:0051100]; negative regulation of cellular organofluorine metabolic process [GO:0090350]; oxidation-reduction process [GO:0055114]; oxidative demethylation [GO:0070989]; steroid metabolic process [GO:0008202]; xenobiotic metabolic process [GO:0006805]
51	A0A087X1C5	CP2D7_HUMAN	CYP2D7	Putative cytochrome P450 2D7 (EC 1.14.14.1)	C461	heme b	Catalytic	1.14.14.1	Cytoplasm, Mitochondrion	Yes		arachidonic acid metabolic process [GO:0019369]; exogenous drug catabolic process [GO:0042738]; xenobiotic metabolic process [GO:0006805]
52	P05181	CP2E1_HUMAN	CYP2E1 CYP2E	Cytochrome P450 2E1 (EC 1.14.13.-) (4-nitrophenol 2-hydroxylase) (EC 1.14.13.n7) (CYPIIE1) (Cytochrome P450-J)	C437	heme b	Catalytic	1.14.13.-; 1.14.13.n7	Endoplasmic reticulum	Yes		benzene metabolic process [GO:0018910]; carbon tetrachloride metabolic process [GO:0018885]; drug metabolic process [GO:0017144]; epoxygenase P450 pathway [GO:0019373]; halogenated hydrocarbon metabolic process [GO:0042197]; heterocycle metabolic process [GO:0046483]; monoterpenoid metabolic process [GO:0016098]; oxidation-reduction process [GO:0055114]; response to drug [GO:0042493]; response to ethanol [GO:0045471]; response to organonitrogen compound [GO:0010243]; response to ozone [GO:0010193]; steroid metabolic process [GO:0008202]; triglyceride metabolic process [GO:0006641]; xenobiotic metabolic process [GO:0006805]
53	P24903	CP2F1_HUMAN	CYP2F1	Cytochrome P450 2F1 (EC 1.14.14.1) (CYPIIF1)	C436	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes		epoxygenase P450 pathway [GO:0019373]; naphthalene metabolic process [GO:0018931]; response to toxic substance [GO:0009636]; trichloroethylene metabolic process [GO:0018979]; xenobiotic metabolic process [GO:0006805]

54	P51589	CP2J2_HUMAN	CYP2J2	Cytochrome P450 2J2 (EC 1.14.14.1) (Arachidonic acid epoxygenase) (CYPIIJ2)	C448	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes		epoxygenase P450 pathway [GO:0019373]; icosanoid metabolic process [GO:0006690]; linoleic acid metabolic process [GO:0043651]; regulation of heart contraction [GO:0008016]; xenobiotic metabolic process [GO:0006805]
55	Q6VVX0	CP2R1_HUMAN	CYP2R1	Vitamin D 25-hydroxylase (EC 1.14.14.24) (Cytochrome P450 2R1)	C448	heme b	Catalytic	1.14.14.24	Endoplasmic reticulum	Yes	DISEASE: Rickets vitamin D-dependent 1B (VDDR1B) [MIM:600081]: A disorder caused by a selective deficiency of the active form of vitamin D (1,25-dihydroxyvitamin D3) and resulting in defective bone mineralization and clinical features of rickets. The patients sera have low calcium concentrations, low phosphate concentrations, elevated alkaline phosphatase activity and low levels of 25-hydroxyvitamin D. {ECO:0000269|PubMed:15128933, ECO:0000269|PubMed:25942481}. Note=The disease is caused by mutations affecting the gene represented in this entry.	response to cesium ion [GO:0010164]; response to ionizing radiation [GO:0010212]; vitamin D metabolic process [GO:0042359]; vitamin metabolic process [GO:0006766]
56	Q96SQ9	CP2S1_HUMAN	CYP2S1 UNQ891/PRO1906	Cytochrome P450 2S1 (EC 1.14.14.1) (CYPIIS1)	C440	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes		epoxygenase P450 pathway [GO:0019373]
57	Q7Z449	CP2U1_HUMAN	CYP2U1	Cytochrome P450 2U1 (EC 1.14.14.1)	C490	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes	DISEASE: Spastic paraplegia 56, autosomal recessive (SPG56) [MIM:615030]: A form of spastic paraplegia, a neurodegenerative disorder characterized by a slow, gradual, progressive weakness and spasticity of the lower limbs. Rate of progression and the severity of symptoms are quite variable. Initial symptoms may include difficulty with balance, weakness and stiffness in the legs, muscle spasms, and dragging the toes when walking. Complicated forms are recognized by additional variable features including spastic quadriparesis, seizures, dementia, amyotrophy, extrapyramidal disturbance, cerebral or cerebellar atrophy, optic atrophy, and peripheral neuropathy, as well as by extra neurological manifestations. In SPG56, upper limbs are often also affected. Some SPG56 patients may have a subclinical axonal neuropathy. {ECO:0000269|PubMed:23176821}. Note=The disease is caused by mutations affecting the gene represented in this entry.	omega-hydroxylase P450 pathway [GO:0097267]
58	Q8TAV3	CP2W1_HUMAN	CYP2W1	Cytochrome P450 2W1 (EC 1.14.14.-) (CYPIIW1)	C433	heme b	Catalytic	1.14.14.-	Endoplasmic reticulum, Cell membrane	Yes		aflatoxin B1 metabolic process [GO:0043390]; epoxygenase P450 pathway [GO:0019373]; xenobiotic metabolic process [GO:0006805]
59	Q9HB55	CP343_HUMAN	CYP3A43	Cytochrome P450 3A43 (EC 1.14.14.1)	C442	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes		
60	Q9NYL5	CP39A_HUMAN	CYP39A1	24-hydroxycholesterol 7-alpha-hydroxylase (EC 1.14.14.26) (Cytochrome P450 39A1) (hCYP39A1) (Oxysterol 7-alpha-hydroxylase)	C414	heme b	Catalytic	1.14.14.26	Endoplasmic reticulum	Yes		bile acid biosynthetic process [GO:0006699]; bile acid catabolic process [GO:0030573]; cholesterol catabolic process [GO:0006707]; digestion [GO:0007586]; sterol metabolic process [GO:0016125]
61	P08684	CP3A4_HUMAN	CYP3A4 CYP3A3	Cytochrome P450 3A4 (EC 1.14.13.-) (1,8-cineole 2-exo-monooxygenase) (EC 1.14.13.157) (Albendazole monooxygenase) (EC 1.14.13.32) (Albendazole sulfoxidase) (CYPIIIA3) (CYPIIIA4) (Cholesterol 25-hydroxylase) (EC 1.14.14.1) (Cytochrome P450 3A3) (Cytochrome P450 HLp) (Cytochrome P450 NF-25) (Cytochrome P450-PCN1) (Nifedipine oxidase) (Quinine 3-monooxygenase) (EC 1.14.13.67) (Taurochenodeoxycholate 6-alpha-hydroxylase) (EC 1.14.13.97)	C442	heme b	Catalytic	1.14.13.-; 1.14.13.157; 1.14.13.32; 1.14.14.1; 1.14.13.67; 1.14.13.97	Endoplasmic reticulum	Yes		alkaloid catabolic process [GO:0009822]; androgen metabolic process [GO:0008209]; drug catabolic process [GO:0042737]; drug metabolic process [GO:0017144]; exogenous drug catabolic process [GO:0042738]; heterocycle metabolic process [GO:0046483]; lipid hydroxylation [GO:0002933]; lipid metabolic process [GO:0006629]; monoterpenoid metabolic process [GO:0016098]; oxidation-reduction process [GO:0055114]; oxidative demethylation [GO:0070989]; steroid catabolic process [GO:0006706]; steroid metabolic process [GO:0008202]; vitamin D metabolic process [GO:0042359]; xenobiotic metabolic process [GO:0006805]
62	P20815	CP3A5_HUMAN	CYP3A5	Cytochrome P450 3A5 (EC 1.14.14.1) (CYPIIIA5) (Cytochrome P450 HLp2) (Cytochrome P450-PCN3)	C441	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes		alkaloid catabolic process [GO:0009822]; drug catabolic process [GO:0042737]; lipid hydroxylation [GO:0002933]; oxidative demethylation [GO:0070989]; steroid metabolic process [GO:0008202]; xenobiotic metabolic process [GO:0006805]
63	P24462	CP3A7_HUMAN	CYP3A7	Cytochrome P450 3A7 (EC 1.14.14.1) (CYPIIIA7) (Cytochrome P450-HFLA)	C442	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes		lipid hydroxylation [GO:0002933]; steroid metabolic process [GO:0008202]; xenobiotic metabolic process [GO:0006805]
64	Q9Y6A2	CP46A_HUMAN	CYP46A1 CYP46	Cholesterol 24-hydroxylase (CH24H) (EC 1.14.14.25) (Cytochrome P450 46A1)	C437	heme b	Catalytic	1.14.14.25	Endoplasmic reticulum	Yes		bile acid biosynthetic process [GO:0006699]; cholesterol catabolic process [GO:0006707]; nervous system development [GO:0007399]; sterol metabolic process [GO:0016125]; xenobiotic metabolic process [GO:0006805]

65	Q02928	CP4AB_HUMAN	CYP4A11 CYP4A2	Cytochrome P450 4A11 (20-hydroxyeicosatetraenoic acid synthase) (20-HETE synthase) (CYP4AII) (CYPIVA11) (Cytochrome P-450HK-omega) (Cytochrome P450HL-omega) (Fatty acid omega-hydroxylase) (Lauric acid omega-hydroxylase) (Long-chain fatty acid omega-monooxygenase) (EC 1.14.13.205)	E321-C457	heme b	Catalytic	1.14.13.205	Endoplasmic reticulum	Yes		arachidonic acid metabolic process [GO:0019369]; epoxygenase P450 pathway [GO:0019373]; fatty acid metabolic process [GO:0006631]; leukotriene metabolic process [GO:0006691]; long-chain fatty acid metabolic process [GO:0001676]; omega-hydroxylase P450 pathway [GO:0097267]; oxidation-reduction process [GO:0055114]; positive regulation of icosanoid secretion [GO:0032305]; pressure natriuresis [GO:0003095]; regulation of lipid metabolic process [GO:0019216]; renal water homeostasis [GO:0003091]; sodium ion homeostasis [GO:0055078]
66	Q5TCH4	CP4AM_HUMAN	CYP4A22	Cytochrome P450 4A22 (CYPIVA22) (Fatty acid omega-hydroxylase) (Lauric acid omega-hydroxylase) (Long-chain fatty acid omega-monooxygenase) (EC 1.14.13.205)	E321-C457	heme b	Catalytic	1.14.13.205	Endoplasmic reticulum	Yes		lipid hydroxylation [GO:0002933]
67	P13584	CP4B1_HUMAN	CYP4B1	Cytochrome P450 4B1 (EC 1.14.14.1) (CYPIVB1) (Cytochrome P450-HP)	E315-C453	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes		biphenyl metabolic process [GO:0018879]; exogenous drug catabolic process [GO:0042738]; fluorene metabolic process [GO:0018917]
68	P78329	CP4F2_HUMAN	CYP4F2	Phylloquinone omega-hydroxylase CYP4F2 (EC 1.14.13.194) (20-hydroxyeicosatetraenoic acid synthase) (20-HETE synthase) (EC 1.14.13.-) (Arachidonic acid omega-hydroxylase) (CYPIVF2) (Cytochrome P450 4F2) (Cytochrome P450-LTB-omega) (Leukotriene-B(4) 20-monooxygenase 1) (Leukotriene-B(4) omega-hydroxylase 1) (EC 1.14.13.30)	E328-C468	heme b	Catalytic	1.14.13.194; 1.14.13.-; 1.14.13.30	Endoplasmic reticulum	Yes	DISEASE: Coumarin resistance (CMRES) [MIM:122700]: A condition characterized by partial or complete resistance to warfarin or other 4-hydroxycoumarin derivatives. These drugs are used as anti-coagulants for the prevention of thromboembolic diseases in subjects with deep vein thrombosis, atrial fibrillation, or mechanical heart valve replacement. Note=Disease susceptibility may be associated with variations affecting the gene represented in this entry. The variant Met-433 is associated with coumarin (the brand name of warfarin) resistance by increasing coumarin maintenance dose in patients on this anti-coagulant therapy. This is probably due to decreased activity of the phylloquinone omega-hydroxylase activity, leading to an increase in hepatic vitamin K levels that warfarin must antagonize (PubMed:24138531). {ECO:0000269|PubMed:24138531}.	arachidonic acid metabolic process [GO:0019369]; blood coagulation [GO:0007596]; drug metabolic process [GO:0017144]; epoxygenase P450 pathway [GO:0019373]; icosanoid metabolic process [GO:0006690]; leukotriene B4 catabolic process [GO:0036101]; leukotriene metabolic process [GO:0006691]; long-chain fatty acid metabolic process [GO:0001676]; menaquinone catabolic process [GO:0042361]; negative regulation of icosanoid secretion [GO:0032304]; omega-hydroxylase P450 pathway [GO:0097267]; oxidation-reduction process [GO:0055114]; phylloquinone catabolic process [GO:0042376]; positive regulation of icosanoid secretion [GO:0032305]; pressure natriuresis [GO:0003095]; regulation of blood pressure [GO:0008217]; renal water homeostasis [GO:0003091]; sodium ion homeostasis [GO:0055078]; very long-chain fatty acid metabolic process [GO:0000038]; vitamin E metabolic process [GO:0042360]; vitamin K catabolic process [GO:0042377]
69	Q08477	CP4F3_HUMAN	CYP4F3 LTB4H	Docosahexaenoic acid omega-hydroxylase CYP4F3 (EC 1.14.13.199) (20-hydroxyeicosatetraenoic acid synthase) (20-HETE synthase) (EC 1.14.13.-) (CYPIVF3) (Cytochrome P450 4F3) (Cytochrome P450-LTB-omega) (Leukotriene-B(4) 20-monooxygenase 2) (Leukotriene-B(4) omega-hydroxylase 2) (EC 1.14.13.30)	E328-C468	heme b	Catalytic	1.14.13.199; 1.14.13.-; 1.14.13.30	Endoplasmic reticulum	Yes		icosanoid metabolic process [GO:0006690]; leukotriene metabolic process [GO:0006691]
70	P98187	CP4F8_HUMAN	CYP4F8	Cytochrome P450 4F8 (EC 1.14.14.1) (CYPIVF8)	C468	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes		icosanoid metabolic process [GO:0006690]; prostaglandin metabolic process [GO:0006693]
71	Q9HBI6	CP4FB_HUMAN	CYP4F11	Phylloquinone omega-hydroxylase CYP4F11 (EC 1.14.13.194) (3-hydroxy fatty acids omega-hydroxylase CYP4F11) (EC 1.14.13.-) (Cytochrome P450 4F11) (CYPIVF11)	E328-C468	heme b	Catalytic	1.14.13.194; 1.14.13.-	Unknown	Yes		blood coagulation [GO:0007596]; fatty acid metabolic process [GO:0006631]; inflammatory response [GO:0006954]; menaquinone catabolic process [GO:0042361]; oxidation-reduction process [GO:0055114]; phylloquinone catabolic process [GO:0042376]; vitamin K catabolic process [GO:0042377]
72	Q9HCS2	CP4FC_HUMAN	CYP4F12 UNQ568/PRO1129	Cytochrome P450 4F12 (EC 1.14.14.1) (CYPIVF12)	C468	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes		arachidonic acid metabolic process [GO:0019369]; drug metabolic process [GO:0017144]; epoxygenase P450 pathway [GO:0019373]; leukotriene B4 catabolic process [GO:0036101]; long-chain fatty acid metabolic process [GO:0001676]; oxidation-reduction process [GO:0055114]; pressure natriuresis [GO:0003095]; renal water homeostasis [GO:0003091]; sodium ion homeostasis [GO:0055078]; very long-chain fatty acid metabolic process [GO:0000038]; vitamin E metabolic process [GO:0042360]

73	Q6NT55	CP4FN_HUMAN	CYP4F22	Cytochrome P450 4F22 (EC 1.14.14.-)	E335-C475	heme b	Catalytic	1.14.14.-	Endoplasmic reticulum	Yes	DISEASE: Ichthyosis, congenital, autosomal recessive 5 (ARCI5) [MIM:604777]: A form of autosomal recessive congenital ichthyosis, a disorder of keratinization with abnormal differentiation and desquamation of the epidermis, resulting in abnormal skin scaling over the whole body. The main skin phenotypes are lamellar ichthyosis (LI) and non-bullous congenital ichthyosiform erythroderma (NCIE), although phenotypic overlap within the same patient or among patients from the same family can occur. Lamellar ichthyosis is a condition often associated with an embedment in a collodion-like membrane at birth; skin scales later develop, covering the entire body surface. Non-bullous congenital ichthyosiform erythroderma characterized by fine whitish scaling on an erythrodermal background; larger brownish scales are present on the buttocks, neck and legs. {ECO:0000269|PubMed:16436457}. Note=The disease is caused by mutations affecting the gene represented in this entry.	icosanoid metabolic process [GO:0006690]
74	Q6ZWL3	CP4V2_HUMAN	CYP4V2	Cytochrome P450 4V2 (EC 1.14.13.-) (Docosahexaenoic acid omega-hydroxylase CYP4V2) (EC 1.14.13.199)	E329-C467	heme b	Catalytic	1.14.13.-; 1.14.13.199	Endoplasmic reticulum	Yes	DISEASE: Bietti crystalline corneoretinal dystrophy (BCD) [MIM:210370]: An autosomal recessive ocular disease characterized by retinal degeneration and marginal corneal dystrophy. Typical features include multiple glistening intraretinal crystals scattered over the fundus, a characteristic degeneration of the retina, and sclerosis of the choroidal vessels, ultimately resulting in progressive night blindness and constriction of the visual field. Most patients have similar crystals at the corneoscleral limbus. Patients develop decreased vision, nyctalopia, and paracentral scotomata between the 2nd and 4th decade of life. Later, they develop peripheral visual field loss and marked visual impairment, usually progressing to legal blindness by the 5th or 6th decade of life. {ECO:0000269|PubMed:15042513, ECO:0000269|PubMed:22772592}. Note=The disease is caused by mutations affecting the gene represented in this entry.	fatty acid omega-oxidation [GO:0010430]; response to stimulus [GO:0050896]; retinoid metabolic process [GO:0001523]; sterol metabolic process [GO:0016125]; visual perception [GO:0007601]
75	Q8N118	CP4X1_HUMAN	CYP4X1 UNQ1929/PRO4404	Cytochrome P450 4X1 (EC 1.14.14.1) (CYPIVX1)	C454	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes		
76	Q86W10	CP4Z1_HUMAN	CYP4Z1 UNQ3060/PRO9882	Cytochrome P450 4Z1 (EC 1.14.14.1) (CYPIVZ1)	C452	heme b	Catalytic	1.14.14.1	Endoplasmic reticulum	Yes		
77	Q8N1L4	CP4Z2_HUMAN	CYP4Z2P	Putative inactive cytochrome P450 family member 4Z2	Unknown	heme b	Catalytic		Unknown	Yes		
78	Q16850	CP51A_HUMAN	CYP51A1 CYP51	Lanosterol 14-alpha demethylase (LDM) (EC 1.14.13.70) (CYPLI) (Cytochrome P450 51A1) (Cytochrome P450-14DM) (Cytochrome P45014DM) (Cytochrome P450LI) (Sterol 14-alpha demethylase)	H447-C449	heme b	Catalytic	1.14.13.70	Endoplasmic reticulum	Yes		cholesterol biosynthetic process [GO:0006695]; cholesterol biosynthetic process via 24,25-dihydrolanosterol [GO:0033488]; regulation of cholesterol biosynthetic process [GO:0045540]; steroid biosynthetic process [GO:0006694]; sterol metabolic process [GO:0016125]
79	P22680	CP7A1_HUMAN	CYP7A1 CYP7	Cholesterol 7-alpha-monooxygenase (EC 1.14.14.23) (CYPVII) (Cholesterol 7-alpha-hydroxylase) (Cytochrome P450 7A1)	C444	heme b	Catalytic	1.14.14.23	Endoplasmic reticulum	Yes		bile acid biosynthetic process [GO:0006699]; cellular response to cholesterol [GO:0071397]; cellular response to glucose stimulus [GO:0071333]; cholesterol catabolic process [GO:0006707]; cholesterol homeostasis [GO:0042632]; regulation of bile acid biosynthetic process [GO:0070857]; regulation of lipid metabolic process [GO:0019216]; sterol metabolic process [GO:0016125]

80	O75881	CP7B1_HUMAN	CYP7B1	25-hydroxycholesterol 7-alpha-hydroxylase (EC 1.14.14.29) (Cytochrome P450 7B1) (Oxysterol 7-alpha-hydroxylase)	C449	heme b	Catalytic	1.14.14.29	Endoplasmic reticulum	Yes	DISEASE: Spastic paraplegia 5A, autosomal recessive (SPG5A) [MIM:270800]: A form of spastic paraplegia, a neurodegenerative disorder characterized by a slow, gradual, progressive weakness and spasticity of the lower limbs. Rate of progression and the severity of symptoms are quite variable. Initial symptoms may include difficulty with balance, weakness and stiffness in the legs, muscle spasms, and dragging the toes when walking. In some forms of the disorder, bladder symptoms (such as incontinence) may appear, or the weakness and stiffness may spread to other parts of the body. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Congenital bile acid synthesis defect 3 (CBAS3) [MIM:613812]: A disorder resulting in severe cholestasis, cirrhosis and liver synthetic failure. Hepatic microsomal oxysterol 7-alpha-hydroxylase activity is undetectable. {ECO:0000269|PubMed:9802883}. Note=The disease is caused by mutations affecting the gene represented in this entry.	B cell chemotaxis [GO:0035754]; bile acid biosynthetic process [GO:0006699]; cholesterol metabolic process [GO:0008203]; negative regulation of intracellular estrogen receptor signaling pathway [GO:0033147]; positive regulation of epithelial cell proliferation [GO:0050679]; prostate gland epithelium morphogenesis [GO:0060740]; sterol metabolic process [GO:0016125]
81	Q9UNU6	CP8B1_HUMAN	CYP8B1 CYP12	7-alpha-hydroxycholest-4-en-3-one 12-alpha-hydroxylase (EC 1.14.18.8) (7-alpha-hydroxy-4-cholesten-3-one 12-alpha-hydroxylase) (CYPVIIIB1) (Cytochrome P450 8B1) (Sterol 12-alpha-hydroxylase)	H120-C440	heme b	Catalytic	1.14.18.8	Endoplasmic reticulum	Yes		bile acid biosynthetic process [GO:0006699]; sterol metabolic process [GO:0016125]
82	P08574	CY1_HUMAN	CYC1	Cytochrome c1, heme protein, mitochondrial (Complex III subunit 4) (Complex III subunit IV) (Cytochrome b-c1 complex subunit 4) (Ubiquinol-cytochrome-c reductase complex cytochrome c1 subunit) (Cytochrome c-1)	H125-M244	heme c	Electron transfer		Mitochondrion	Yes	DISEASE: Mitochondrial complex III deficiency, nuclear 6 (MC3DN6) [MIM:615453]: An autosomal recessive disorder caused by mitochondrial dysfunction. It is characterized by onset in early childhood of episodic acute lactic acidosis, ketoacidosis, and insulin-responsive hyperglycemia, usually associated with infection. Laboratory studies show decreased activity of mitochondrial complex III. Psychomotor development is normal. {ECO:0000269|PubMed:23910460}. Note=The disease is caused by mutations affecting the gene represented in this entry.	mitochondrial ATP synthesis coupled proton transport [GO:0042776]; mitochondrial electron transport, ubiquinol to cytochrome c [GO:0006122]; response to glucagon [GO:0033762]
83	P13498	CY24A_HUMAN	CYBA	Cytochrome b-245 light chain (Cytochrome b(558) alpha chain) (Cytochrome b558 subunit alpha) (Neutrophil cytochrome b 22 kDa polypeptide) (Superoxide-generating NADPH oxidase light chain subunit) (p22 phagocyte B-cytochrome) (p22-phox) (p22phox)	H94	heme b	Electron transfer		Cell membrane	Yes	DISEASE: Granulomatous disease, chronic, cytochrome-b-negative, autosomal recessive (ARCGD) [MIM:233690]: A disorder characterized by the inability of neutrophils and phagocytes to kill microbes that they have ingested. Patients suffer from life-threatening bacterial/fungal infections. Note=The disease is caused by mutations affecting the gene represented in this entry.	cell redox homeostasis [GO:0045454]; cellular response to angiotensin [GO:1904385]; cellular response to gamma radiation [GO:0071480]; cellular response to glucose stimulus [GO:0071333]; cellular response to L-glutamine [GO:1904845]; cellular response to mechanical stimulus [GO:0071260]; cellular response to organic cyclic compound [GO:0071407]; cellular response to oxidative stress [GO:0034599]; cellular response to tumor necrosis factor [GO:0071356]; cytochrome complex assembly [GO:0017004]; hydrogen peroxide biosynthetic process [GO:0050665]; inflammatory response [GO:0006954]; positive regulation of endothelial cell proliferation [GO:0001938]; positive regulation of interleukin-6 production [GO:0032755]; positive regulation of mucus secretion [GO:0070257]; positive regulation of reactive oxygen species biosynthetic process [GO:1903428]; positive regulation of smooth muscle cell proliferation [GO:0048661]; positive regulation of superoxide anion generation [GO:0032930]; positive regulation of toll-like receptor 2 signaling pathway [GO:0034137]; positive regulation of tumor necrosis factor production [GO:0032760]; regulation of release of sequestered calcium ion into cytosol [GO:0051279]; respiratory burst [GO:0045730]; response to drug [GO:0042493]; response to hypoxia [GO:0001666]; superoxide metabolic process [GO:0006801]; vascular endothelial growth factor receptor signaling pathway [GO:0048010]

84	P04839	CY24B_HUMAN	CYBB NOX2	Cytochrome b-245 heavy chain (EC 1.-.-.-) (CGD91-phox) (Cytochrome b(558) subunit beta) (Cytochrome b558 subunit beta) (Heme-binding membrane glycoprotein gp91phox) (NADPH oxidase 2) (Neutrophil cytochrome b 91 kDa polypeptide) (Superoxide-generating NADPH oxidase heavy chain subunit) (gp91-1) (gp91-phox) (p22 phagocyte B-cytochrome)	H101-H115; H209-H222	heme b	Electron transfer	1.-.-.-	Cell membrane	Yes	DISEASE: Granulomatous disease, chronic, X-linked (CGD) [MIM:306400]: A disorder characterized by the inability of neutrophils and phagocytes to kill microbes that they have ingested. Patients suffer from life-threatening bacterial/fungal infections. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Immunodeficiency 34 (IMD34) [MIM:300645]: A form of Mendelian susceptibility to mycobacterial disease, a rare condition characterized by predisposition to illness caused by moderately virulent mycobacterial species, such as Bacillus Calmette-Guerin (BCG) vaccine, environmental non-tuberculous mycobacteria, and by the more virulent Mycobacterium tuberculosis. Other microorganisms rarely cause severe clinical disease in individuals with susceptibility to mycobacterial infections, with the exception of Salmonella which infects less than 50% of these individuals. {ECO:0000269|PubMed:21278736}. Note=Disease susceptibility is associated with variations affecting the gene represented in this entry.	antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent [GO:0002479]; cell redox homeostasis [GO:0045454]; cellular response to cadmium ion [GO:0071276]; cellular response to ethanol [GO:0071361]; cellular response to L-glutamine [GO:1904845]; cellular response to oxidative stress [GO:0034599]; electron transport chain [GO:0022900]; hydrogen peroxide biosynthetic process [GO:0050665]; hypoxia-inducible factor-1alpha signaling pathway [GO:0097411]; inflammatory response [GO:0006954]; innate immune response [GO:0045087]; neutrophil degranulation [GO:0043312]; oxidation-reduction process [GO:0055114]; positive regulation of angiogenesis [GO:0045766]; positive regulation of tumor necrosis factor biosynthetic process [GO:0042535]; respiratory burst [GO:0045730]; response to aldosterone [GO:1904044]; response to angiotensin [GO:1990776]; response to drug [GO:0042493]; response to nutrient [GO:0007584]; superoxide anion generation [GO:0042554]; superoxide metabolic process [GO:0006801]; vascular endothelial growth factor receptor signaling pathway [GO:0048010]
85	P49447	CY561_HUMAN	CYB561	Cytochrome b561 (Cytochrome b-561)	H53-H121; H87-H160	heme b	Electron transfer		Unknown	Yes		electron transport chain [GO:0022900]; oxidation-reduction process [GO:0055114]
86	Q8NBI2	CYAC3_HUMAN	CYB561A3 CYBASC3 LCYTB PSEC0259	Cytochrome b ascorbate-dependent protein 3 (EC 1.-.-.-) (Cytochrome b561 family member A3) (Lysosomal cytochrome b) (LCytb)	H47-H117; H83-H156	heme b	Electron transfer	1.-.-.-	Endosome	Yes		oxidation-reduction process [GO:0055114]
87	P00156	CYB_HUMAN	MT-CYB COB CYTB MTCYB	Cytochrome b (Complex III subunit 3) (Complex III subunit III) (Cytochrome b-c1 complex subunit 3) (Ubiquinol-cytochrome-c reductase complex cytochrome b subunit)	H83-H182; H97-H196	heme b	Electron transfer		Mitochondrion	Yes	DISEASE: Note=Defects in MT-CYB are a rare cause of mitochondrial dysfunction underlying different myopathies. They include mitochondrial encephalomyopathy, hypertrophic cardiomyopathy (HCM), and sporadic mitochondrial myopathy (MM). In mitochondrial myopathy, exercise intolerance is the predominant symptom. Additional features include lactic acidosis, muscle weakness and/or myoglobinuria. Defects in MTCYB are also found in cases of exercise intolerance accompanied by deafness, mental retardation, retinitis pigmentosa, cataract, growth retardation, epilepsy (multisystem disorder). {ECO:0000269|PubMed:11047755, ECO:0000269|PubMed:11601507}.; DISEASE: Cardiomyopathy, infantile histiocytoid (CMIH) [MIM:500000]: A heart disease characterized by the presence of pale granular foamy histiocyte-like cells within the myocardium. It usually affects children younger than 2 years of age, with a clear predominance of females over males. Infants present with dysrhythmia or cardiac arrest. The clinical course is usually fulminant, sometimes simulating sudden infant death syndrome. {ECO:0000269|PubMed:10960495}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Leber hereditary optic neuropathy (LHON) [MIM:535000]: A maternally inherited disease resulting in acute or subacute loss of central vision, due to optic nerve dysfunction. Cardiac conduction defects and neurological defects have also been described in some patients. LHON results from primary mitochondrial DNA mutations affecting the respiratory chain complexes. {ECO:0000269|PubMed:1732158}. Note=The disease is caused by mutations affecting distinct genetic loci, including the gene represented in this entry.	animal organ regeneration [GO:0031100]; hyperosmotic salinity response [GO:0042538]; mitochondrial electron transport, ubiquinol to cytochrome c [GO:0006122]; response to cadmium ion [GO:0046686]; response to calcium ion [GO:0051592]; response to cobalamin [GO:0033590]; response to copper ion [GO:0046688]; response to drug [GO:0042493]; response to ethanol [GO:0045471]; response to glucagon [GO:0033762]; response to heat [GO:0009408]; response to hyperoxia [GO:0055093]; response to hypoxia [GO:0001666]; response to mercury ion [GO:0046689]; response to toxic substance [GO:0009636]
88	P00167	CYB5_HUMAN	CYB5A CYB5	Cytochrome b5 (Microsomal cytochrome b5 type A) (MCB5)	H44-H68	heme b	Electron transfer		Cytoplasm, Endoplasmic reticulum	Yes	DISEASE: Methemoglobinemia CYB5A-related (METHB-CYB5A) [MIM:250790]: A form of methemoglobinemia, a hematologic disease characterized by the presence of excessive amounts of methemoglobin in blood cells, resulting in decreased oxygen carrying capacity of the blood, cyanosis and hypoxia. {ECO:0000269|PubMed:8168836}. Note=The disease is caused by mutations affecting the gene represented in this entry.	L-ascorbic acid metabolic process [GO:0019852]; response to cadmium ion [GO:0046686]

89	O43169	CYB5B_HUMAN	CYB5B CYB5M OMB5	Cytochrome b5 type B (Cytochrome b5 outer mitochondrial membrane isoform)	H55-H79; H96	heme b	Electron transfer		Mitochondrion	Yes		oxidation-reduction process [GO:0055114]; xenobiotic metabolic process [GO:0006805]
90	Q53TN4	CYBR1_HUMAN	CYBRD1 DCYTB FRRS3	Cytochrome b reductase 1 (EC 1.-.-.-) (Duodenal cytochrome b) (Ferric-chelate reductase 3)	H50-H120; H86-H159	heme b	Electron transfer	1.-.-.-	Unknown	Yes		cellular iron ion homeostasis [GO:0006879]; response to iron ion [GO:0010039]
91	P99999	CYC_HUMAN	CYCS CYC	Cytochrome c	H19-M81	heme c	Electron transfer		Mitochondrion	Yes	DISEASE: Thrombocytopenia 4 (THC4) [MIM:612004]: Thrombocytopenia is defined by a decrease in the number of platelets in circulating blood, resulting in the potential for increased bleeding and decreased ability for clotting. {ECO:0000269|PubMed:18345000}. Note=The disease is caused by mutations affecting the gene represented in this entry.	activation of cysteine-type endopeptidase activity involved in apoptotic process by cytochrome c [GO:0008635]; cellular respiration [GO:0045333]; cellular response to oxidative stress [GO:0034599]; intrinsic apoptotic signaling pathway [GO:0097193]; mitochondrial electron transport, cytochrome c to oxygen [GO:0006123]; mitochondrial electron transport, ubiquinol to cytochrome c [GO:0006122]; mitochondrion organization [GO:0007005]; protein dephosphorylation [GO:0006470]
92	Q8WWM9	CYGB_HUMAN	CYGB STAP	Cytoglobin (Histoglobin) (HGb) (Stellate cell activation-associated protein)	H81-H113	heme b	Oxygen storage/transport		Cytoplasm	No		fatty acid oxidation [GO:0019395]; negative regulation of collagen biosynthetic process [GO:0032966]; negative regulation of fibroblast migration [GO:0010764]; negative regulation of hepatic stellate cell activation [GO:2000490]; oxygen transport [GO:0015671]; regulation of nitric-oxide synthase activity [GO:0050999]; response to hypoxia [GO:0001666]; response to oxidative stress [GO:0006979]
93	Q8WYQ5	DGCR8_HUMAN	DGCR8 C22orf12 DGCRK6 LP4941	Microprocessor complex subunit DGCR8 (DiGeorge syndrome critical region 8)	C352	heme b	Structural or Regulatory		Nucleus	No		miRNA metabolic process [GO:0010586]; primary miRNA processing [GO:0031053]; regulation of stem cell proliferation [GO:0072091]; RNA phosphodiester bond hydrolysis, endonucleolytic [GO:0090502]
94	O14521	DHSD_HUMAN	SDHD SDH4	Succinate dehydrogenase [ubiquinone] cytochrome b small subunit, mitochondrial (CybS) (CII-4) (QPs3) (Succinate dehydrogenase complex subunit D) (Succinate-ubiquinone oxidoreductase cytochrome b small subunit) (Succinate-ubiquinone reductase membrane anchor subunit)	H102	heme b	Electron transfer		Mitochondrion	Yes	DISEASE: Paragangliomas 1 (PGL1) [MIM:168000]: A neural crest tumor usually derived from the chromoreceptor tissue of a paraganglion. DISEASE: Pheochromocytoma (PCC) [MIM:171300]: A catecholamine-producing tumor of chromaffin tissue of the adrenal medulla or sympathetic paraganglia. DISEASE: Intestinal carcinoid tumor (ICT) [MIM:114900]: A yellow, well-differentiated, circumscribed tumor that arises from enterochromaffin cells in the small intestine or, less frequently, in other parts of the gastrointestinal tract.; DISEASE: Paraganglioma and gastric stromal sarcoma (PGGSS) [MIM:606864]: Gastrointestinal stromal tumors may be sporadic or inherited in an autosomal dominant manner, alone or as a component of a syndrome associated with other tumors, such as in the context of neurofibromatosis type 1 (NF1). DISEASE: Cowden syndrome 3 (CWS3) [MIM:615106]: A form of Cowden syndrome, a hamartomatous polyposis syndrome with age-related penetrance. DISEASE: Mitochondrial complex II deficiency (MT-C2D) [MIM:252011]: A disorder of the mitochondrial respiratory chain with heterogeneous clinical manifestations. Clinical features include psychomotor regression in infants, poor growth with lack of speech development, severe spastic quadriplegia, dystonia, progressive leukoencephalopathy, muscle weakness, exercise intolerance, cardiomyopathy. Some patients manifest Leigh syndrome or Kearns-Sayre syndrome	mitochondrial electron transport, succinate to ubiquinone [GO:0006121]; tricarboxylic acid cycle [GO:0006099]
95	Q9NRD9	DUOX1_HUMAN	DUOX1 DUOX LNOX1 THOX1	Dual oxidase 1 (EC 1.11.1.-) (EC 1.6.3.1) (Large NOX 1) (Long NOX 1) (NADPH thyroid oxidase 1) (Thyroid oxidase 1)	H770-H1225-H1238	heme b	Catalytic	1.11.1.-; 1.6.3.1	Cell membrane	Yes		cuticle development [GO:0042335]; cytokine-mediated signaling pathway [GO:0019221]; hormone biosynthetic process [GO:0042446]; hydrogen peroxide biosynthetic process [GO:0050665]; hydrogen peroxide catabolic process [GO:0042744]; oxidation-reduction process [GO:0055114]; response to cAMP [GO:0051591]; response to oxidative stress [GO:0006979]; superoxide anion generation [GO:0042554]; thyroid hormone generation [GO:0006590]

96	Q9NRD8	DUOX2_HUMAN	DUOX2 LNOX2 THOX2	Dual oxidase 2 (EC 1.11.1.-) (EC 1.6.3.1) (Large NOX 2) (Long NOX 2) (NADH/NADPH thyroid oxidase p138-tox) (NADPH oxidase/peroxidase DUOX2) (NADPH thyroid oxidase 2) (Thyroid oxidase 2) (p138 thyroid oxidase)	H774-H1222-H1235	heme b	Catalytic	1.11.1.-; 1.6.3.1	Cell membrane, Cell membrane	Yes	DISEASE: Thyroid dyshormonogenesis 6 (TDH6) [MIM:607200]: A disorder due to a defective conversion of accumulated iodide to organically bound iodine. The iodide organification defect can be partial or complete. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Note=Defects in DUOX2 may play a role in the pathogenesis of very early onset inflammatory bowel disease (VEOIBD), a chronic, relapsing inflammation of the gastrointestinal tract with a complex etiology diagnosed before 6 years of age. VEOIBD is subdivided into Crohn disease and ulcerative colitis phenotypes. Crohn disease may affect any part of the gastrointestinal tract from the mouth to the anus, but the phenotype of children with onset of Crohn disease occurring younger than the age of 10 is predominantly colonic, with a lower risk of ileal disease. Bowel inflammation is transmural and discontinuous; it may contain granulomas or be associated with intestinal or perianal fistulas. In contrast, in ulcerative colitis, the inflammation is continuous and limited to rectal and colonic mucosal layers; fistulas and granulomas are not observed. Both diseases include extraintestinal inflammation of the skin, eyes, or joints. {ECO:0000269|PubMed:26301257}.	adenohypophysis morphogenesis [GO:0048855]; bone mineralization [GO:0030282]; cuticle development [GO:0042335]; cytokine-mediated signaling pathway [GO:0019221]; fertilization [GO:0009566]; hormone biosynthetic process [GO:0042446]; hydrogen peroxide biosynthetic process [GO:0050665]; hydrogen peroxide catabolic process [GO:0042744]; inner ear development [GO:0048839]; multicellular organism growth [GO:0035264]; oxidation-reduction process [GO:0055114]; response to cAMP [GO:0051591]; response to oxidative stress [GO:0006979]; response to virus [GO:0009615]; thyroid gland development [GO:0030878]; thyroid hormone generation [GO:0006590]
97	Q9BQI3	E2AK1_HUMAN	EIF2AK1 HRI KIAA1369 PRO1362	Eukaryotic translation initiation factor 2-alpha kinase 1 (EC 2.7.11.1) (Heme-controlled repressor) (HCR) (Heme-regulated eukaryotic initiation factor eIF-2-alpha kinase) (Heme-regulated inhibitor) (Hemin-sensitive initiation factor 2-alpha kinase)	Unknown	heme b	Unknown	2.7.11.1	Cytoplasm	No		acute inflammatory response [GO:0002526]; iron ion homeostasis [GO:0055072]; macrophage differentiation [GO:0030225]; negative regulation of cell proliferation [GO:0008285]; negative regulation of hemoglobin biosynthetic process [GO:0046986]; negative regulation of translational initiation by iron [GO:0045993]; phagocytosis [GO:0006909]; protein autophosphorylation [GO:0046777]; protoporphyrinogen IX metabolic process [GO:0046501]; regulation of eIF2 alpha phosphorylation by heme [GO:0010999]; response to external stimulus [GO:0009605]; response to stress [GO:0006950]
98	Q6ZMW3	EMAL6_HUMAN	EML6 EML5L	Echinoderm microtubule-associated protein-like 6 (EMAP-6) (Echinoderm microtubule-associated protein-like 5-like)	Unknown	heme d1	Catalytic		Cytoplasm	No		
99	Q7L5A8	FA2H_HUMAN	FA2H FAAH	Fatty acid 2-hydroxylase (EC 1.-.-.-) (Fatty acid alpha-hydroxylase)	H43-H69	heme b	Electron transfer	1.-.-.-	Endoplasmic reticulum	Yes	DISEASE: Spastic paraplegia 35, autosomal recessive (SPG35) [MIM:612319]: A form of spastic paraplegia, a neurodegenerative disorder characterized by a slow, gradual, progressive weakness and spasticity of the lower limbs. Rate of progression and the severity of symptoms are quite variable. Initial symptoms may include difficulty with balance, weakness and stiffness in the legs, muscle spasms, and dragging the toes when walking. In some forms of the disorder, bladder symptoms (such as incontinence) may appear, or the weakness and stiffness may spread to other parts of the body. SPG35 is a complicated form characterized by childhood onset of gait difficulties. It has a rapid progression and many patients become wheelchair-bound as young adults. Patients manifest cognitive decline associated with leukodystrophy. Other variable neurologic features, such as dystonia, optic atrophy, and seizures may also occur. {ECO:0000269|PubMed:19068277, ECO:0000269|PubMed:20104589, ECO:0000269|PubMed:20853438}. Note=The disease is caused by mutations affecting the gene represented in this entry.	central nervous system myelin maintenance [GO:0032286]; fatty acid biosynthetic process [GO:0006633]; fatty acid metabolic process [GO:0006631]; lipid modification [GO:0030258]; peripheral nervous system myelin maintenance [GO:0032287]; regulation of cell proliferation [GO:0042127]; regulation of hair cycle [GO:0042634]; sebaceous gland cell differentiation [GO:0001949]; sphingolipid biosynthetic process [GO:0030148]
100	O60427	FADS1_HUMAN	FADS1 FADSD5	Fatty acid desaturase 1 (EC 1.14.19.-) (Delta(5) fatty acid desaturase) (D5D) (Delta(5) desaturase) (Delta-5 desaturase)	H52-H75; H138-H183	heme b	Electron transfer	1.14.19.-	Endoplasmic reticulum, Mitochondrion	Yes		alpha-linolenic acid metabolic process [GO:0036109]; cell-cell signaling [GO:0007267]; cellular response to starvation [GO:0009267]; icosanoid biosynthetic process [GO:0046456]; linoleic acid metabolic process [GO:0043651]; phospholipid biosynthetic process [GO:0008654]; regulation of cell differentiation [GO:0045595]; regulation of lipid metabolic process [GO:0019216]; regulation of transcription, DNA-templated [GO:0006355]; unsaturated fatty acid biosynthetic process [GO:0006636]
101	O95864	FADS2_HUMAN	FADS2	Fatty acid desaturase 2 (EC 1.14.19.3) (Acyl-CoA 6-desaturase) (Delta(6) fatty acid desaturase) (D6D) (Delta(6) desaturase) (Delta-6 desaturase)	H53-H76; H184	heme b	Electron transfer	1.14.19.3	Endoplasmic reticulum	Yes		alpha-linolenic acid metabolic process [GO:0036109]; linoleic acid metabolic process [GO:0043651]; unsaturated fatty acid biosynthetic process [GO:0006636]

102	Q9Y5Q0	FADS3_HUMAN	FADS3 CYB5RP	Fatty acid desaturase 3 (EC 1.14.19.-) (Cytochrome b5-related protein)	H55; H78-H186	heme b	Electron transfer	1.14.19.-	Endoplasmic reticulum	Yes		unsaturated fatty acid biosynthetic process [GO:0006636]
103	P02771	FETA_HUMAN	AFP HPAFP	Alpha-fetoprotein (Alpha-1-fetoprotein) (Alpha-fetoglobulin)	Y185-Y377	heme b	Substrate - transport		Extracellular space	No	DISEASE: Alpha-fetoprotein deficiency (AFPD) [MIM:615969]: A benign condition characterized by undetectable AFP levels in the amniotic fluid. Affected individuals are asymptomatic and present normal development. {ECO:0000269|PubMed:15280901, ECO:0000269|PubMed:18854864}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Alpha-fetoprotein, hereditary persistence (HPAFP) [MIM:615970]: A benign autosomal dominant condition characterized by continued expression of alpha-fetoprotein in adult life. {ECO:0000269|PubMed:7684942}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cellular protein metabolic process [GO:0044267]; ovulation from ovarian follicle [GO:0001542]; post-translational protein modification [GO:0043687]; progesterone metabolic process [GO:0042448]; SMAD protein signal transduction [GO:0060395]; transport [GO:0006810]
104	Q9Y5Y0	FLVC1_HUMAN	FLVCR1 FLVCR	Feline leukemia virus subgroup C receptor-related protein 1 (Feline leukemia virus subgroup C receptor) (hFLVCR)	Unknown	heme b, Precursor	Substrate - transport		Cell membrane	Yes	DISEASE: Posterior column ataxia with retinitis pigmentosa (PCARP) [MIM:609033]: A neurodegenerative syndrome beginning in infancy with areflexia and retinitis pigmentosa. Nyctalopia (night blindness) and peripheral visual field loss are usually evident during late childhood or teenage years, with subsequent progressive constriction of the visual fields and loss of central retinal function over time. A sensory ataxia caused by degeneration of the posterior columns of the spinal cord results in a loss of proprioceptive sensation that is clinically evident in the second decade of life and gradually progresses. Scoliosis, camptodactyly, achalasia, gastrointestinal dysmotility, and a sensory peripheral neuropathy are variable features of the disease. Affected individuals have no clinical or radiological evidence of cerebral or cerebellar involvement. DISEASE: Note=Defects in FLVCR1 are a cause of a sensory neuropathy resulting in pain insensitivity. Patients have decreased sensing of pain, temperature and touch. Self-injury, ulcers and amputations are commonly observed in affected individuals. {ECO:0000269|PubMed:27923065}.	blood vessel development [GO:0001568]; cellular iron ion homeostasis [GO:0006879]; embryonic digit morphogenesis [GO:0042733]; embryonic skeletal system morphogenesis [GO:0048704]; erythrocyte differentiation [GO:0030218]; erythrocyte maturation [GO:0043249]; head morphogenesis [GO:0060323]; heme export [GO:0097037]; heme transport [GO:0015886]; in utero embryonic development [GO:0001701]; mitochondrial transport [GO:0006839]; multicellular organism development [GO:0007275]; multicellular organism growth [GO:0035264]; regulation of organ growth [GO:0046620]; spleen development [GO:0048536]; transmembrane transport [GO:0055085]; transport [GO:0006810]
105	Q9UPI3	FLVC2_HUMAN	FLVCR2 C14orf58	Feline leukemia virus subgroup C receptor-related protein 2 (Calcium-chelate transporter) (CCT)	Unknown	heme b, Precursor	Substrate - transport		Cell membrane	Yes	DISEASE: Proliferative vasculopathy and hydranencephaly-hydrocephaly syndrome (PVHH) [MIM:225790]: A rare prenatally lethal disorder characterized by hydranencephaly, a distinctive glomerular vasculopathy in the central nervous system and retina, and diffuse ischemic lesions of the brain stem, basal ganglia, and spinal cord with calcifications. Hydranencephaly is a condition where the greater portions of the cerebral hemispheres and corpus striatum are replaced by cerebrospinal fluid and glial tissue. {ECO:0000269|PubMed:20206334, ECO:0000269|PubMed:20518025, ECO:0000269|PubMed:20690116}. Note=The disease is caused by mutations affecting the gene represented in this entry.	transmembrane transport [GO:0055085]
106	Q86XX4	FRAS1_HUMAN	FRAS1 KIAA1500	Extracellular matrix protein FRAS1	H1799; H1945; H2080-H3301	heme b, heme c	Unknown		Cell membrane	Yes	DISEASE: Fraser syndrome 1 (FRASRS1) [MIM:219000]: A form of Fraser syndrome, an autosomal recessive disorder characterized by cryptophthalmos, cutaneous syndactyly, and urogenital abnormalities including renal agenesis or hypoplasia. Additional features include abnormalities of the larynx, ear malformations, and facial abnormalities. {ECO:0000269|PubMed:12766769, ECO:0000269|PubMed:23473829}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cell communication [GO:0007154]; embryonic limb morphogenesis [GO:0030326]; metanephros morphogenesis [GO:0003338]; morphogenesis of an epithelium [GO:0002009]; palate development [GO:0060021]; protein transport [GO:0015031]; skin development [GO:0043588]
107	Q6ZNA5	FRRS1_HUMAN	FRRS1 SDFR2 SDR2	Ferric-chelate reductase 1 (EC 1.-.-.-) (Stromal cell-derived receptor 2) (SDR-2)	H373-H414; H446-H482	heme b	Electron transfer	1.-.-.-	Unknown	Yes		

108	Q9P0K9	FRS1L_HUMAN	FRRS1L C9orf4	DOMON domain-containing protein FRRS1L (Brain protein CG-6) (Ferric-chelate reductase 1-like protein)	M205	heme b	Unknown		Cell membrane	Yes	DISEASE: Epileptic encephalopathy, early infantile, 37 (EIEE37) [MIM:616981]: A form of epileptic encephalopathy, a heterogeneous group of severe childhood onset epilepsies characterized by refractory seizures, neurodevelopmental impairment, and poor prognosis. Development is normal prior to seizure onset, after which cognitive and motor delays become apparent. EIEE37 is an autosomal recessive, severe form manifesting in the first years of life. Affected individuals show hyperkinetic movement disorder with choreoathetosis, spasticity, rigidity, mental retardation, absent speech, and impaired volitional movements. {ECO:0000269|PubMed:27236917, ECO:0000269|PubMed:27239025}. Note=The disease is caused by mutations affecting the gene represented in this entry.	regulation of glutamate receptor signaling pathway [GO:1900449]
109	A8MWK0	FS2P1_HUMAN	FADS2P1	Putative fatty acid desaturase 2-like protein FADS2P1 (Fatty acid desaturase 2 pseudogene 1)	H90-H113	heme b	Electron transfer		Endoplasmic reticulum	Yes		unsaturated fatty acid biosynthetic process [GO:0006636]
110	P33402	GCYA2_HUMAN	GUCY1A2 GUC1A2 GUCSA2	Guanylate cyclase soluble subunit alpha-2 (GCS-alpha-2) (EC 4.6.1.2)	H480	heme b	Regulatory - Gaseus sensor which activate catalysis	4.6.1.2	Cytoplasm	No		cGMP biosynthetic process [GO:0006182]; intracellular signal transduction [GO:0035556]; positive regulation of cGMP biosynthetic process [GO:0030828]; signal transduction [GO:0007165]
111	Q02108	GCYA3_HUMAN	GUCY1A3 GUC1A3 GUCSA3 GUCY1A1	Guanylate cyclase soluble subunit alpha-3 (GCS-alpha-3) (EC 4.6.1.2) (GCS-alpha-1) (Soluble guanylate cyclase large subunit)	Unknown	heme b	Regulatory - Gaseus sensor which activate catalysis	4.6.1.2	Cytoplasm	No	DISEASE: Moyamoya disease 6 with achalasia (MYMY6) [MIM:615750]: A form of Moyamoya disease, a progressive cerebral angiopathy characterized by bilateral intracranial carotid artery stenosis and telangiectatic vessels in the region of the basal ganglia. The abnormal vessels resemble a 'puff of smoke' (moyamoya) on cerebral angiogram. Affected individuals can develop transient ischemic attacks and/or cerebral infarction, and rupture of the collateral vessels can cause intracranial hemorrhage. Hemiplegia of sudden onset and epileptic seizures constitute the prevailing presentation in childhood, while subarachnoid bleeding occurs more frequently in adults. MYMY6 is characterized by severe cerebral angiopathy and onset of severe achalasia in infancy or early childhood. {ECO:0000269|PubMed:24581742}. Note=The disease is caused by mutations affecting the gene represented in this entry.	blood circulation [GO:0008015]; cGMP biosynthetic process [GO:0006182]; nitric oxide mediated signal transduction [GO:0007263]; positive regulation of cGMP biosynthetic process [GO:0030828]; regulation of blood pressure [GO:0008217]; relaxation of vascular smooth muscle [GO:0060087]; response to defense-related host nitric oxide production [GO:0052565]
112	Q02153	GCYB1_HUMAN	GUCY1B3 GUC1B3 GUCSB3 GUCY1B1	Guanylate cyclase soluble subunit beta-1 (GCS-beta-1) (EC 4.6.1.2) (Guanylate cyclase soluble subunit beta-3) (GCS-beta-3) (Soluble guanylate cyclase small subunit)	H105	heme b	Regulatory - Gaseus sensor which activate catalysis	4.6.1.2	Cytoplasm	No		blood circulation [GO:0008015]; cellular response to nitric oxide [GO:0071732]; cGMP biosynthetic process [GO:0006182]; nitric oxide-cGMP-mediated signaling pathway [GO:0038060]; nitric oxide mediated signal transduction [GO:0007263]
113	O75343	GCYB2_HUMAN	GUCY1B2	Guanylate cyclase soluble subunit beta-2 (GCS-beta-2) (EC 4.6.1.2)	H26	heme b	Regulatory - Gaseus sensor which activate catalysis	4.6.1.2	Cytoplasm	No		cGMP biosynthetic process [GO:0006182]; intracellular signal transduction [GO:0035556]; signal transduction [GO:0007165]
114	P69905	HBA_HUMAN	HBA1; HBA2	Hemoglobin subunit alpha (Alpha-globin) (Hemoglobin alpha chain)	H59-H88	heme b	Oxygen storage/transport		Unknown	No	DISEASE: Heinz body anemias (HEIBAN) [MIM:140700]: Form of non-spherocytic hemolytic anemia of Dacie type 1. ; DISEASE: Alpha-thalassemia (A-THAL) [MIM:604131]: A form of thalassemia. Thalassemias are common monogenic diseases occurring mostly in Mediterranean and Southeast Asian populations. DISEASE: Note=Alpha(0)-thalassemia is associated with non-immune hydrops fetalis, a generalized edema of the fetus with fluid accumulation in the body cavities due to non-immune causes. Non-immune hydrops fetalis is not a diagnosis in itself but a symptom, a feature of many genetic disorders, and the end-stage of a wide variety of disorders.; DISEASE: Hemoglobin H disease (HBH) [MIM:613978]: A form of alpha-thalassemia due to the loss of three alpha genes. This results in high levels of a tetramer of four beta chains (hemoglobin H), causing a severe and life-threatening anemia. Untreated, most patients die in childhood or early adolescence. {ECO:0000269|PubMed:10569720}. Note=The disease is caused by mutations affecting the gene represented in this entry.	bicarbonate transport [GO:0015701]; cellular oxidant detoxification [GO:0098869]; hydrogen peroxide catabolic process [GO:0042744]; oxygen transport [GO:0015671]; positive regulation of cell death [GO:0010942]; protein heterooligomerization [GO:0051291]; receptor-mediated endocytosis [GO:0006898]; response to hydrogen peroxide [GO:0042542]
115	P09105	HBAT_HUMAN	HBQ1	Hemoglobin subunit theta-1 (Hemoglobin theta-1 chain) (Theta-1-globin)	H59-H88	heme b	Oxygen storage/transport		Unknown	No		oxygen transport [GO:0015671]
116	P02008	HBAZ_HUMAN	HBZ HBZ2	Hemoglobin subunit zeta (HBAZ) (Hemoglobin zeta chain) (Zeta-globin)	H59-H88	heme b	Oxygen storage/transport		Unknown	No		erythrocyte maturation [GO:0043249]; negative regulation of transcription from RNA polymerase II promoter [GO:0000122]

117	P68871	HBB_HUMAN	HBB	Hemoglobin subunit beta (Beta-globin) (Hemoglobin beta chain) [Cleaved into: LVV-hemorphin-7; Spinorphin]	H64-H93	heme b	Oxygen storage/transport		Unknown	No	DISEASE: Heinz body anemias (HEIBAN) [MIM:140700]: Form of non-spherocytic hemolytic anemia of Dacie type 1. After splenectomy, which has little benefit, basophilic inclusions called Heinz bodies are demonstrable in the erythrocytes. DISEASE: Beta-thalassemia (B-THAL) [MIM:613985]: A form of thalassemia. Thalassemias are common monogenic diseases occurring mostly in Mediterranean and Southeast Asian populations. The hallmark of beta-thalassemia is an imbalance in globin-chain production in the adult HbA molecule. DISEASE: Sickle cell anemia (SKCA) [MIM:603903]: Characterized by abnormally shaped red cells resulting in chronic anemia and periodic episodes of pain, serious infections and damage to vital organs. Normal red blood cells are round and flexible and flow easily through blood vessels, but in sickle cell anemia, the abnormal hemoglobin (called Hb S) causes red blood cells to become stiff. They are C-shaped and resembles a sickle. These stiffer red blood cells can led to microvascular occlusion thus cutting off the blood supply to nearby tissues. DISEASE: Beta-thalassemia, dominant, inclusion body type (B-THALIB) [MIM:603902]: An autosomal dominant form of beta thalassemia characterized by moderate anemia, lifelong jaundice, cholelithiasis and splenomegaly, marked morphologic changes in the red cells, erythroid hyperplasia of the bone marrow with increased numbers of multinucleate red cell precursors, and the presence of large inclusion bodies in the normoblasts, both in the marrow and in the peripheral blood after splenectomy.	bicarbonate transport [GO:0015701]; blood coagulation [GO:0007596]; cellular oxidant detoxification [GO:0098869]; hydrogen peroxide catabolic process [GO:0042744]; neutrophil degranulation [GO:0043312]; nitric oxide transport [GO:0030185]; oxygen transport [GO:0015671]; platelet aggregation [GO:0070527]; positive regulation of cell death [GO:0010942]; positive regulation of nitric oxide biosynthetic process [GO:0045429]; protein heterooligomerization [GO:0051291]; receptor-mediated endocytosis [GO:0006898]; regulation of blood pressure [GO:0008217]; regulation of blood vessel size [GO:0050880]; renal absorption [GO:0070293]; response to hydrogen peroxide [GO:0042542]
118	P02042	HBD_HUMAN	HBD	Hemoglobin subunit delta (Delta-globin) (Hemoglobin delta chain)	H64-H93	heme b	Oxygen storage/transport		Unknown	No		blood coagulation [GO:0007596]
119	P02100	HBE_HUMAN	HBE1 HBE	Hemoglobin subunit epsilon (Epsilon-globin) (Hemoglobin epsilon chain)	H64-H93	heme b	Oxygen storage/transport		Unknown	No		blood coagulation [GO:0007596]; protein heterooligomerization [GO:0051291]; response to organic cyclic compound [GO:0014070]
120	P69891	HBG1_HUMAN	HBG1 PRO2979	Hemoglobin subunit gamma-1 (Gamma-1-globin) (Hb F Agamma) (Hemoglobin gamma-1 chain) (Hemoglobin gamma-A chain)	H64-H93	heme b	Oxygen storage/transport		Unknown	No		blood coagulation [GO:0007596]
121	P69892	HBG2_HUMAN	HBG2	Hemoglobin subunit gamma-2 (Gamma-2-globin) (Hb F Ggamma) (Hemoglobin gamma-2 chain) (Hemoglobin gamma-G chain)	H64-H93	heme b	Oxygen storage/transport		Unknown	No	DISEASE: Cyanosis transient neonatal (TNCY) [MIM:613977]: A disorder characterized by cyanosis in the fetus and neonate, due to a defect in the fetal hemoglobin chain which has reduced affinity for oxygen. Some patients develop anemia resulting from increased destruction of red cells containing abnormal or unstable hemoglobin. The cyanosis resolves spontaneously by 5 to 6 months of age or earlier, as the adult beta-globin chain is produced and replaces the fetal gamma-globin chain. Note=The disease is caused by mutations affecting the gene represented in this entry.	blood coagulation [GO:0007596]
122	Q6B0K9	HBM_HUMAN	HBM HBAP2	Hemoglobin subunit mu (Hemoglobin mu chain) (Mu-globin)	H58-H87	heme b	Oxygen storage/transport		Unknown	No		
123	Q9NRV9	HEBP1_HUMAN	HEBP1 HBP	Heme-binding protein 1 (p22HBP)	Unknown	Variuos types	Substrate - transport		Cytoplasm	No		circadian rhythm [GO:0007623]; G-protein coupled receptor signaling pathway [GO:0007186]
124	Q9Y5Z4	HEBP2_HUMAN	HEBP2 C6orf34 SOUL	Heme-binding protein 2 (Placental protein 23) (PP23) (Protein SOUL)	Unknown	Variuos types	Substrate - transport		Cytoplasm, Mitochondrion	No		negative regulation of mitochondrial membrane potential [GO:0010917]; neutrophil degranulation [GO:0043312]; positive regulation of mitochondrial membrane permeability [GO:0035794]; positive regulation of necrotic cell death [GO:0010940]

125	P22830	HEMH_HUMAN	FECH	Ferrochelatase, mitochondrial (EC 4.99.1.1) (Heme synthase) (Protoheme ferro-lyase)	H263	heme b	Substrate - Biosynthesis	4.99.1.1	Mitochondrion	Yes	DISEASE: Erythropoietic protoporphyria (EPP) [MIM:177000]: A form of porphyria. Porphyrias are inherited defects in the biosynthesis of heme, resulting in the accumulation and increased excretion of porphyrins or porphyrin precursors. They are classified as erythropoietic or hepatic, depending on whether the enzyme deficiency occurs in red blood cells or in the liver. Erythropoietic protoporphyria is marked by excessive protoporphyrin in erythrocytes, plasma, liver and feces, and by widely varying photosensitive skin changes ranging from a burning or pruritic sensation to erythema, edema and wheals. Note=The disease is caused by mutations affecting the gene represented in this entry.	cellular response to dexamethasone stimulus [GO:0071549]; generation of precursor metabolites and energy [GO:0006091]; heme biosynthetic process [GO:0006783]; protoporphyrinogen IX metabolic process [GO:0046501]; response to arsenic-containing substance [GO:0046685]; response to drug [GO:0042493]; response to ethanol [GO:0045471]; response to insecticide [GO:0017085]; response to lead ion [GO:0010288]; response to light stimulus [GO:0009416]; response to methylmercury [GO:0051597]; response to platinum ion [GO:0070541]
126	P02790	HEMO_HUMAN	HPX	Hemopexin (Beta-1B-glycoprotein)	H293	heme b	Substrate - degradation		Extracellular space	No		cellular iron ion homeostasis [GO:0006879]; heme metabolic process [GO:0042168]; heme transport [GO:0015886]; hemoglobin metabolic process [GO:0020027]; positive regulation of humoral immune response mediated by circulating immunoglobulin [GO:0002925]; positive regulation of immunoglobulin production [GO:0002639]; positive regulation of interferon-gamma-mediated signaling pathway [GO:0060335]; positive regulation of tyrosine phosphorylation of STAT protein [GO:0042531]; receptor-mediated endocytosis [GO:0006898]; viral process [GO:0016032]
127	O95714	HERC2_HUMAN	HERC2	E3 ubiquitin-protein ligase HERC2 (EC 2.3.2.26) (HECT domain and RCC1-like domain-containing protein 2) (HECT-type E3 ubiquitin transferase HERC2)	Unknown	heme b	Electron transfer	2.3.2.26	Cytoplasm, Nucleus	No	DISEASE: Mental retardation, autosomal recessive 38 (MRT38) [MIM:615516]: A disorder characterized by significantly below average general intellectual functioning associated with impairments in adaptive behavior and manifested during the developmental period. MRT38 is characterized by global developmental delay affecting motor, speech, adaptive, and social development. Patients manifest autistic features, aggression, self-injury, impulsivity, and distractibility. {ECO:0000269|PubMed:23065719}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cellular response to DNA damage stimulus [GO:0006974]; double-strand break repair via nonhomologous end joining [GO:0006303]; intracellular protein transport [GO:0006886]; proteasome-mediated ubiquitin-dependent protein catabolic process [GO:0043161]; protein ubiquitination [GO:0016567]; spermatogenesis [GO:0007283]
128	P09601	HMOX1_HUMAN	HMOX1 HO HO1	Heme oxygenase 1 (HO-1) (EC 1.14.14.18)	H25-E29	heme b	Substrate - degradation	1.14.14.18	Cytoplasm, Endoplasmic reticulum	Yes	DISEASE: Heme oxygenase 1 deficiency (HMOX1D) [MIM:614034]: A disease characterized by impaired stress hematopoiesis, resulting in marked erythrocyte fragmentation and intravascular hemolysis, coagulation abnormalities, endothelial damage, and iron deposition in renal and hepatic tissues. Clinical features include persistent hemolytic anemia, asplenia, nephritis, generalized erythematous rash, growth retardation and hepatomegaly. {ECO:0000269|PubMed:9884342}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cell death [GO:0008219]; cellular iron ion homeostasis [GO:0006879]; cellular response to arsenic-containing substance [GO:0071243]; cellular response to cadmium ion [GO:0071276]; cellular response to cisplatin [GO:0072719]; cellular response to heat [GO:0034605]; cellular response to hypoxia [GO:0071456]; cellular response to nutrient [GO:0031670]; endothelial cell proliferation [GO:0001935]; heme catabolic process [GO:0042167]; heme oxidation [GO:0006788]; intracellular signal transduction [GO:0035556]; iron ion homeostasis [GO:0055072]; liver regeneration [GO:0097421]; low-density lipoprotein particle clearance [GO:0034383]; negative regulation of DNA binding [GO:0043392]; negative regulation of leukocyte migration [GO:0002686]; negative regulation of macroautophagy [GO:0016242]; negative regulation of mast cell cytokine production [GO:0032764]; negative regulation of mast cell degranulation [GO:0043305]; negative regulation of muscle cell apoptotic process [GO:0010656]; negative regulation of neuron apoptotic process [GO:0043524]; negative regulation of smooth muscle cell proliferation [GO:0048662]; response to estrogen [GO:0043627]; response to hydrogen peroxide [GO:0042542]; response to nicotine [GO:0035094]; response to oxidative stress [GO:0006979]; small GTPase mediated signal transduction [GO:0007264]; smooth muscle hyperplasia [GO:0014806]; wound healing involved in inflammatory response [GO:0002246]
129	P30519	HMOX2_HUMAN	HMOX2 HO2	Heme oxygenase 2 (HO-2) (EC 1.14.14.18)	H45-E49	heme b	Substrate - degradation	1.14.14.18	Endoplasmic reticulum	No		cellular iron ion homeostasis [GO:0006879]; heme catabolic process [GO:0042167]; heme oxidation [GO:0006788]; iron ion homeostasis [GO:0055072]; neutrophil degranulation [GO:0043312]; response to hypoxia [GO:0001666]; response to oxidative stress [GO:0006979]

130	P04196	HRG_HUMAN	HRG	Histidine-rich glycoprotein (Histidine-proline-rich glycoprotein) (HPRG)	Unknown	Unknown	Substrate - transport		Extracellular space	No	DISEASE: Thrombophilia due to histidine-rich glycoprotein deficiency (THPH11) [MIM:613116]: A hemostatic disorder characterized by a tendency to thrombosis. {ECO:0000269|PubMed:11057869, ECO:0000269|PubMed:9414276}. Note=The disease is caused by mutations affecting the gene represented in this entry.	angiogenesis [GO:0001525]; antimicrobial humoral immune response mediated by antimicrobial peptide [GO:0061844]; chemotaxis [GO:0006935]; cytolysis in other organism [GO:0051715]; defense response to fungus [GO:0050832]; fibrinolysis [GO:0042730]; heme transport [GO:0015886]; negative regulation of angiogenesis [GO:0016525]; negative regulation of blood vessel endothelial cell migration [GO:0043537]; negative regulation of cell adhesion [GO:0007162]; negative regulation of cell adhesion mediated by integrin [GO:0033629]; negative regulation of cell growth [GO:0030308]; negative regulation of cell proliferation [GO:0008285]; negative regulation of endothelial cell chemotaxis [GO:2001027]; negative regulation of fibrinolysis [GO:0051918]; negative regulation of lamellipodium assembly [GO:0010593]; negative regulation of vascular endothelial growth factor signaling pathway [GO:1900747]; platelet activation [GO:0030168]; platelet degranulation [GO:0002576]; positive regulation of apoptotic process [GO:0043065]; positive regulation of blood vessel remodeling [GO:2000504]; positive regulation of focal adhesion assembly [GO:0051894]; positive regulation of immune response to tumor cell [GO:0002839]; regulation of actin cytoskeleton organization [GO:0032956]; regulation of blood coagulation [GO:0030193]; regulation of gene expression [GO:0010468]; regulation of peptidyl-tyrosine phosphorylation [GO:0050730]; regulation of platelet activation [GO:0010543]; regulation of protein complex assembly [GO:0043254]
131	Q6P1K1	HRG1_HUMAN	SLC48A1 HRG1	Heme transporter HRG1 (Heme-responsive gene 1 protein homolog) (HRG-1) (hHRG-1) (Solute carrier family 48 member 1)	Unknown	heme b	Substrate - transport		Endosome	Yes		heme transport [GO:0015886]
132	P14902	I23O1_HUMAN	IDO1 IDO INDO	Indoleamine 2,3-dioxygenase 1 (IDO-1) (EC 1.13.11.52) (Indoleamine-pyrrole 2,3-dioxygenase)	H346	heme b	Catalytic	1.13.11.52	Cytoplasm	No		cytokine production involved in inflammatory response [GO:0002534]; female pregnancy [GO:0007565]; immune system process [GO:0002376]; kynurenic acid biosynthetic process [GO:0034276]; multicellular organismal response to stress [GO:0033555]; negative regulation of interleukin-10 production [GO:0032693]; negative regulation of T cell apoptotic process [GO:0070233]; negative regulation of T cell proliferation [GO:0042130]; positive regulation of chronic inflammatory response [GO:0002678]; positive regulation of interleukin-12 production [GO:0032735]; positive regulation of T cell apoptotic process [GO:0070234]; positive regulation of T cell tolerance induction [GO:0002666]; positive regulation of type 2 immune response [GO:0002830]; response to lipopolysaccharide [GO:0032496]; swimming behavior [GO:0036269]; tryptophan catabolic process [GO:0006569]; tryptophan catabolic process to kynurenine [GO:0019441]
133	Q6ZQW0	I23O2_HUMAN	IDO2 INDOL1	Indoleamine 2,3-dioxygenase 2 (IDO-2) (EC 1.13.11.-) (Indoleamine 2,3-dioxygenase-like protein 1) (Indoleamine-pyrrole 2,3-dioxygenase-like protein 1)	H360	heme b	Catalytic	1.13.11.-	Unknown	No		immune system process [GO:0002376]; tryptophan catabolic process [GO:0006569]; tryptophan catabolic process to kynurenine [GO:0019441]

134	O60674	JAK2_HUMAN	JAK2	Tyrosine-protein kinase JAK2 (EC 2.7.10.2) (Janus kinase 2) (JAK-2)	Unknown	heme b	Regulatory - catalysis	2.7.10.2	Cytoplasm, Nucleus	Yes	DISEASE: Note=Chromosomal aberrations involving JAK2 are found in both chronic and acute forms of eosinophilic, lymphoblastic and myeloid leukemia. Translocation t(8;9)(p22;p24) with PCM1 links the protein kinase domain of JAK2 to the major portion of PCM1. Translocation t(9;12)(p24;p13) with ETV6.; DISEASE: Budd-Chiari syndrome (BDCHS) [MIM:600880]: A syndrome caused by obstruction of hepatic venous outflow involving either the hepatic veins or the terminal segment of the inferior vena cava. DISEASE: Polycythemia vera (PV) [MIM:263300]: A myeloproliferative disorder characterized by abnormal proliferation of all hematopoietic bone marrow elements, erythroid hyperplasia, an absolute increase in total blood volume, but also by myeloid leukocytosis, thrombocytosis and splenomegaly. DISEASE: Thrombocythemia 3 (THCYT3) [MIM:614521]: A myeloproliferative disorder characterized by excessive platelet production, resulting in increased numbers of circulating platelets. It can be associated with spontaneous hemorrhages and thrombotic episodes. DISEASE: Myelofibrosis (MYELOF) [MIM:254450]: A disorder characterized by replacement of the bone marrow by fibrous tissue, occurring in association with a myeloproliferative disorder. Clinical manifestations may include anemia, pallor, splenomegaly, hypermetabolic state, petechiae, ecchymosis, bleeding, lymphadenopathy, hepatomegaly, portal hypertension. DISEASE: Leukemia, acute myelogenous (AML) [MIM:601626]: A subtype of acute leukemia, a cancer of the white blood cells. AML is a malignant disease of bone marrow characterized by maturational arrest of hematopoietic precursors at an early stage of development.	actin filament polymerization [GO:0030041]; activation of cysteine-type endopeptidase activity involved in apoptotic process [GO:0006919]; apoptotic process [GO:0006915]; axon regeneration [GO:0031103]; blood coagulation [GO:0007596]; cell differentiation [GO:0030154]; extrinsic apoptotic signaling pathway [GO:0097191]; intrinsic apoptotic signaling pathway in response to oxidative stress [GO:0008631]; negative regulation of cardiac muscle cell apoptotic process [GO:0010667]; negative regulation of cell-cell adhesion [GO:0022408]; negative regulation of cell proliferation [GO:0008285]; negative regulation of DNA binding [GO:0043392]; negative regulation of heart contraction [GO:0045822]; negative regulation of neuron apoptotic process [GO:0043524]; positive regulation of DNA binding [GO:0043388]; positive regulation of epithelial cell apoptotic process [GO:1904037]; positive regulation of growth factor dependent skeletal muscle satellite cell proliferation [GO:1902728]; positive regulation of growth hormone receptor signaling pathway [GO:0060399]; positive regulation of inflammatory response [GO:0050729]; positive regulation of insulin secretion process [GO:0051770]; positive regulation of peptidyl-tyrosine phosphorylation [GO:0050731]; regulation of apoptotic process [GO:0042981]; regulation of cell proliferation [GO:0042127]; regulation of inflammatory response [GO:0050727]; regulation of interferon-gamma-mediated signaling pathway [GO:0060334]; response to antibiotic [GO:0046677]; response to hydroperoxide [GO:0033194]; response to interleukin-12 [GO:0070671]; response to lipopolysaccharide [GO:0032496]; response to tumor necrosis factor [GO:0034612]; signal transduction [GO:0007165]; STAT protein import into nucleus [GO:0007262]; tumor necrosis factor-mediated signaling pathway [GO:0033209]; tyrosine phosphorylation of STAT protein [GO:0007260]
135	P03952	KLKB1_HUMAN	KLKB1 KLK3	Plasma kallikrein (EC 3.4.21.34) (Fletcher factor) (Kininogenin) (Plasma prekallikrein) (PKK) [Cleaved into: Plasma kallikrein heavy chain; Plasma kallikrein light chain]	C66	heme b	Catalytic	3.4.21.34	Extracellular space	No	DISEASE: Prekallikrein deficiency (PKK deficiency) [MIM:612423]: This disorder is a blood coagulation defect. {ECO:0000269|PubMed:14652634, ECO:0000269|PubMed:17598838}. Note=The disease is caused by mutations affecting the gene represented in this entry.	blood coagulation, intrinsic pathway [GO:0007597]; extracellular matrix disassembly [GO:0022617]; Factor XII activation [GO:0002542]; fibrinolysis [GO:0042730]; plasminogen activation [GO:0031639]; positive regulation of fibrinolysis [GO:0051919]; proteolysis [GO:0006508]; zymogen activation [GO:0031638]
136	Q6UVY6	MOXD1_HUMAN	MOXD1 MOX UNQ2493/PRO5780	DBH-like monooxygenase protein 1 (EC 1.14.17.-) (Monooxygenase X)	M70	heme b	Unknown	1.14.17.-	Endoplasmic reticulum	Yes		dopamine catabolic process [GO:0042420]; norepinephrine biosynthetic process [GO:0042421]; octopamine biosynthetic process [GO:0006589]
137	P02144	MYG_HUMAN	MB	Myoglobin	H65-H94	heme b	Oxygen storage/transport		Unknown	No		brown fat cell differentiation [GO:0050873]; enucleate erythrocyte differentiation [GO:0043353]; heart development [GO:0007507]; oxygen transport [GO:0015671]; response to hormone [GO:0009725]; response to hydrogen peroxide [GO:0042542]; response to hypoxia [GO:0001666]; slow-twitch skeletal muscle fiber contraction [GO:0031444]
138	Q7L1T6	NB5R4_HUMAN	CYB5R4 NCB5OR	Cytochrome b5 reductase 4 (EC 1.6.2.2) (Flavohemoprotein b5/b5R) (b5+b5R) (N-terminal cytochrome b5 and cytochrome b5 oxidoreductase domain-containing protein) (cb5/cb5R)	H89-H112	heme b	Electron transfer	1.6.2.2	Endoplasmic reticulum	No		bicarbonate transport [GO:0015701]; cell development [GO:0048468]; detection of oxygen [GO:0003032]; generation of precursor metabolites and energy [GO:0006091]; glucose homeostasis [GO:0042593]; insulin secretion [GO:0030073]; oxidation-reduction process [GO:0055114]; response to antibiotic [GO:0046677]; superoxide metabolic process [GO:0006801]
139	Q9UMX5	NENF_HUMAN	NENF CIR2 SPUF	Neudesin (Cell immortalization-related protein 2) (Neuron-derived neurotrophic factor) (Protein GIG47) (Secreted protein of unknown function) (SPUF protein)	Y88	heme b	Unknown		Extracellular space, Extracellular space	No		negative regulation of appetite [GO:0032099]; positive regulation of MAPK cascade [GO:0043410]
140	Q8WUJ1	NEUFC_HUMAN	CYB5D2	Neuferricin (Cytochrome b5 domain-containing protein 2)	Y79	heme b	Electron transfer		Extracellular space	No		positive regulation of neuron differentiation [GO:0045666]
141	Q9NPG2	NGB_HUMAN	NGB	Neuroglobin	H64-H96	heme b	Oxygen storage/transport		Cytoplasm, Mitochondrion	No		apoptotic process [GO:0006915]; oxygen transport [GO:0015671]

142	P29475	NOS1_HUMAN	NOS1	Nitric oxide synthase, brain (EC 1.14.13.39) (Constitutive NOS) (NC-NOS) (NOS type I) (Neuronal NOS) (N-NOS) (nNOS) (Peptidyl-cysteine S-nitrosylase NOS1) (bNOS)	C420	heme b	Catalytic	1.14.13.39	Cell membrane	Yes		arginine catabolic process [GO:0006527]; cell redox homeostasis [GO:0045454]; cellular response to growth factor stimulus [GO:0071363]; exogenous drug catabolic process [GO:0042738]; multicellular organismal response to stress [GO:0033555]; myoblast fusion [GO:0007520]; negative regulation of adrenergic receptor signaling pathway involved in heart process [GO:1901205]; negative regulation of blood pressure [GO:0045776]; negative regulation of calcium ion transport [GO:0051926]; negative regulation of calcium ion transport into cytosol [GO:0010523]; negative regulation of hydrolase activity [GO:0051346]; negative regulation of potassium ion transport [GO:0043267]; negative regulation of serotonin uptake [GO:0051612]; neurotransmitter biosynthetic process [GO:0042136]; nitric oxide biosynthetic process [GO:0006809]; nitric oxide mediated signal transduction [GO:0007263]; peptidyl-cysteine S-nitrosylation [GO:0018119]; positive regulation of adrenergic receptor signaling pathway involved in heart process [GO:1901206]; positive regulation of guanylate cyclase activity [GO:0031284]; positive regulation of histone acetylation [GO:0035066]; positive regulation of sodium ion transmembrane transport [GO:1902307]; positive regulation of the force of heart contraction [GO:0098735]; positive regulation of transcription, DNA-templated [GO:0045893]; positive regulation of transcription from RNA polymerase II promoter [GO:0045944]; regulation of calcium ion transmembrane transport via high voltage-gated calcium channel [GO:1902514]; regulation of cardiac conduction [GO:1903779]; regulation of cardiac muscle contraction [GO:0055117]; regulation of neurogenesis [GO:0050767]; regulation of ryanodine-sensitive calcium-release channel activity [GO:0060314]; regulation of sodium ion transport [GO:0002028]; response to heat [GO:0009408]; response to hypoxia [GO:0001666]; retrograde trans-synaptic signaling by nitric oxide [GO:0098924]; striated muscle contraction [GO:0006941]; vasodilation [GO:0042311]
143	P35228	NOS2_HUMAN	NOS2 NOS2A	Nitric oxide synthase, inducible (EC 1.14.13.39) (Hepatocyte NOS) (HEP-NOS) (Inducible NO synthase) (Inducible NOS) (iNOS) (NOS type II) (Peptidyl-cysteine S-nitrosylase NOS2)	C200	heme b	Catalytic	1.14.13.39	Unknown	No		arginine catabolic process [GO:0006527]; cell redox homeostasis [GO:0045454]; cellular response to drug [GO:0035690]; cellular response to interferon-gamma [GO:0071346]; cellular response to lipopolysaccharide [GO:0071222]; circadian rhythm [GO:0007623]; defense response to bacterium [GO:0042742]; defense response to Gram-negative bacterium [GO:0050829]; innate immune response in mucosa [GO:0002227]; interleukin-6 secretion [GO:0072604]; interleukin-8 secretion [GO:0072606]; negative regulation of blood pressure [GO:0045776]; negative regulation of gene expression [GO:0010629]; negative regulation of protein catabolic process [GO:0042177]; nitric oxide biosynthetic process [GO:0006809]; nitric oxide mediated signal transduction [GO:0007263]; peptidyl-cysteine S-nitrosylation [GO:0018119]; positive regulation of blood vessel diameter [GO:0097755]; positive regulation of guanylate cyclase activity [GO:0031284]; positive regulation of killing of cells of other organism [GO:0051712]; positive regulation of leukocyte mediated cytotoxicity [GO:0001912]; prostaglandin secretion [GO:0032310]; regulation of cell proliferation [GO:0042127]; regulation of cellular respiration [GO:0043457]; regulation of cytokine production involved in inflammatory response [GO:1900015]; regulation of insulin secretion [GO:0050796]; response to bacterium [GO:0009617]; response to hypoxia [GO:0001666]; superoxide metabolic process [GO:0006801]

144	P29474	NOS3_HUMAN	NOS3	Nitric oxide synthase, endothelial (EC 1.14.13.39) (Constitutive NOS) (cNOS) (EC-NOS) (Endothelial NOS) (eNOS) (NOS type III) (NOSIII)	C184	heme b	Catalytic	1.14.13.39	Cytoplasm, Golgi apparatus, Cell membrane	Yes	DISEASE: Note=Variation in NOS3 seem to be associated with susceptibility to coronary spasm. {ECO:0000269|PubMed:11740345, ECO:0000269|PubMed:9737779}.	angiogenesis [GO:0001525]; arginine catabolic process [GO:0006527]; blood vessel remodeling [GO:0001974]; cell redox homeostasis [GO:0045454]; endothelial cell migration [GO:0043542]; in utero embryonic development [GO:0001701]; lipopolysaccharide-mediated signaling pathway [GO:0031663]; lung development [GO:0030324]; mitochondrion organization [GO:0007005]; negative regulation of blood pressure [GO:0045776]; negative regulation of calcium ion transport [GO:0051926]; negative regulation of cell proliferation [GO:0008285]; negative regulation of potassium ion transport [GO:0043267]; nitric oxide biosynthetic process [GO:0006809]; nitric oxide mediated signal transduction [GO:0007263]; ovulation from ovarian follicle [GO:0001542]; positive regulation of angiogenesis [GO:0045766]; positive regulation of blood vessel diameter [GO:0097755]; positive regulation of guanylate cyclase activity [GO:0031284]; regulation of blood pressure [GO:0008217]; regulation of blood vessel size [GO:0050880]; regulation of nitric-oxide synthase activity [GO:0050999]; regulation of sodium ion transport [GO:0002028]; regulation of systemic arterial blood pressure by endothelin [GO:0003100]; regulation of the force of heart contraction by chemical signal [GO:0003057]; removal of superoxide radicals [GO:0019430]; response to fluid shear stress [GO:0034405]; response to heat [GO:0009408]; smooth muscle hyperplasia [GO:0014806]; vasodilation [GO:0042311]
145	Q9Y5S8	NOX1_HUMAN	NOX1 MOX1 NOH1	NADPH oxidase 1 (NOX-1) (EC 1.-.-.-) (Mitogenic oxidase 1) (MOX-1) (NADH/NADPH mitogenic oxidase subunit P65-MOX) (NOH-1)	H101-H115; H209-H221	heme b	Electron transfer	1.-.-.-	Cell membrane	Yes	DISEASE: Note=Defects in NOX1 may play a role in the pathogenesis of very early onset inflammatory bowel disease (VEOIBD), a chronic, relapsing inflammation of the gastrointestinal tract with a complex etiology diagnosed before 6 years of age. VEOIBD is subdivided into Crohn disease and ulcerative colitis phenotypes. Crohn disease may affect any part of the gastrointestinal tract from the mouth to the anus, but the phenotype of children with onset of Crohn disease occurring younger than the age of 10 is predominantly colonic, with a lower risk of ileal disease. Bowel inflammation is transmural and discontinuous; it may contain granulomas or be associated with intestinal or perianal fistulas. In contrast, in ulcerative colitis, the inflammation is continuous and limited to rectal and colonic mucosal layers; fistulas and granulomas are not observed. Both diseases include extraintestinal inflammation of the skin, eyes, or joints. {ECO:0000269|PubMed:26301257}.	angiogenesis [GO:0001525]; cell migration [GO:0016477]; cellular response to hyperoxia [GO:0071455]; cellular stress response to acidic pH [GO:1990451]; extracellular matrix organization [GO:0030198]; hydrogen peroxide metabolic process [GO:0042743]; inflammatory response [GO:0006954]; intracellular pH elevation [GO:0051454]; NADP metabolic process [GO:0006739]; oxidation-reduction process [GO:0055114]; oxygen metabolic process [GO:0072592]; positive regulation of cell proliferation [GO:0008284]; positive regulation of integrin biosynthetic process [GO:0045726]; positive regulation of JNK cascade [GO:0046330]; positive regulation of oxidative stress-induced intrinsic apoptotic signaling pathway [GO:1902177]; positive regulation of smooth muscle cell proliferation [GO:0048661]; positive regulation of vascular endothelial growth factor production [GO:0010575]; proton transport [GO:0015992]; regulation of blood pressure [GO:0008217]; regulation of systemic arterial blood pressure by renin-angiotensin [GO:0003081]; respiratory burst [GO:0045730]; signal transduction [GO:0007165]; superoxide anion generation [GO:0042554]; superoxide metabolic process [GO:0006801]

146	Q9NPH5	NOX4_HUMAN	NOX4 RENOX	NADPH oxidase 4 (EC 1.6.3.-) (Kidney oxidase-1) (KOX-1) (Kidney superoxide-producing NADPH oxidase) (Renal NAD(P)H-oxidase)	H105-H119; H194-H207	heme b	Electron transfer	1.6.3.-	Endoplasmic reticulum, Cell membrane, Nucleus, Cell membrane	Yes		bone resorption [GO:0045453]; cardiac muscle cell differentiation [GO:0055007]; cell aging [GO:0007569]; cell morphogenesis [GO:0000902]; cellular response to cAMP [GO:0071320]; cellular response to gamma radiation [GO:0071480]; cellular response to glucose stimulus [GO:0071333]; cellular response to oxidative stress [GO:0034599]; cellular response to transforming growth factor beta stimulus [GO:0071560]; gene expression [GO:0010467]; homocysteine metabolic process [GO:0050667]; inflammatory response [GO:0006954]; negative regulation of cell proliferation [GO:0008285]; oxidation-reduction process [GO:0055114]; positive regulation of apoptotic process [GO:0043065]; positive regulation of DNA biosynthetic process [GO:2000573]; positive regulation of ERK1 and ERK2 cascade [GO:0070374]; positive regulation of MAP kinase activity [GO:0043406]; positive regulation of protein kinase B signaling [GO:0051897]; positive regulation of reactive oxygen species metabolic process [GO:2000379]; positive regulation of smooth muscle cell migration [GO:0014911]; positive regulation of stress fiber assembly [GO:0051496]; reactive oxygen species metabolic process [GO:0072593]; response to hypoxia [GO:0001666]; superoxide anion generation [GO:0042554]; superoxide metabolic process [GO:0006801]
147	Q96PH1	NOX5_HUMAN	NOX5	NADPH oxidase 5 (EC 1.6.3.-)	H314-H328; H402-H415	heme b	Electron transfer	1.6.3.-	Unknown	Yes		angiogenesis [GO:0001525]; apoptotic process [GO:0006915]; cell proliferation [GO:0008283]; cellular response to oxidative stress [GO:0034599]; cytokine secretion [GO:0050663]; cytokinesis [GO:0000910]; endothelial cell proliferation [GO:0001935]; oxidation-reduction process [GO:0055114]; positive regulation of reactive oxygen species metabolic process [GO:2000379]; regulation of fusion of sperm to egg plasma membrane [GO:0043012]; regulation of proton transport [GO:0010155]; superoxide anion generation [GO:0042554]
148	Q99743	NPAS2_HUMAN	NPAS2 BHLHE9 MOP4 PASD4	Neuronal PAS domain-containing protein 2 (Neuronal PAS2) (Basic-helix-loop-helix-PAS protein MOP4) (Class E basic helix-loop-helix protein 9) (bHLHe9) (Member of PAS protein 4) (PAS domain-containing protein 4)	H119-H171	heme b	Regulatory - trascription		Nucleus	No		cellular response to DNA damage stimulus [GO:0006974]; central nervous system development [GO:0007417]; circadian regulation of gene expression [GO:0032922]; circadian rhythm [GO:0007623]; negative regulation of cell death [GO:0060548]; positive regulation of DNA repair [GO:0045739]; positive regulation of transcription, DNA-templated [GO:0045893]; positive regulation of transcription from RNA polymerase II promoter [GO:0045944]; regulation of lipid metabolic process [GO:0019216]; regulation of response to DNA damage stimulus [GO:2001020]; response to redox state [GO:0051775]; transcription, DNA-templated [GO:0006351]

149	P20393	NR1D1_HUMAN	NR1D1 EAR1 HREV THRAL	Nuclear receptor subfamily 1 group D member 1 (Rev-erbA-alpha) (V-erbA-related protein 1) (EAR-1)	H602	heme b	Substrate - Regulatory/Sensor		Cytoplasm, Nucleus	No		cell differentiation [GO:0030154]; cellular response to lipopolysaccharide [GO:0071222]; circadian regulation of gene expression [GO:0032922]; circadian rhythm [GO:0007623]; circadian temperature homeostasis [GO:0060086]; glycogen biosynthetic process [GO:0005978]; negative regulation of receptor biosynthetic process [GO:0010871]; negative regulation of toll-like receptor 4 signaling pathway [GO:0034144]; negative regulation of transcription, DNA-templated [GO:0045892]; negative regulation of transcription from RNA polymerase II promoter [GO:0000122]; positive regulation of bile acid biosynthetic process [GO:0070859]; positive regulation of transcription, DNA-templated [GO:0045893]; proteasomal protein catabolic process [GO:0010498]; regulation of cholesterol homeostasis [GO:2000188]; regulation of circadian rhythm [GO:0042752]; regulation of fat cell differentiation [GO:0045598]; regulation of gluconeogenesis by regulation of transcription from RNA polymerase II promoter [GO:0035947]; regulation of insulin secretion involved in cellular response to glucose stimulus [GO:0061178]; regulation of lipid metabolic process [GO:0019216]; regulation of type B pancreatic cell proliferation [GO:0061469]; response to leptin [GO:0044321]; transcription initiation from RNA polymerase II promoter [GO:0006367]
150	Q14995	NR1D2_HUMAN	NR1D2	Nuclear receptor subfamily 1 group D member 2 (Orphan nuclear hormone receptor BD73) (Rev-erb alpha-related receptor) (RVR) (Rev-erb-beta) (V-erbA-related protein 1-related) (EAR-1R)	H568	heme b	Regulatory - trascription		Nucleus	No		lipid homeostasis [GO:0055088]; negative regulation of transcription, DNA-templated [GO:0045892]; positive regulation of transcription, DNA-templated [GO:0045893]; regulation of circadian rhythm [GO:0042752]; regulation of energy homeostasis [GO:2000505]; regulation of inflammatory response [GO:0050727]; regulation of lipid metabolic process [GO:0019216]; regulation of skeletal muscle cell differentiation [GO:2001014]; regulation of transcription, DNA-templated [GO:0006355]; rhythmic process [GO:0048511]; transcription initiation from RNA polymerase II promoter [GO:0006367]
151	Q96NT5	PCFT_HUMAN	SLC46A1 HCP1 PCFT	Proton-coupled folate transporter (G21) (Heme carrier protein 1) (PCFT/HCP1) (Solute carrier family 46 member 1)	Unknown	Unknown	Substrate - transport		Cytoplasm, Cell membrane	Yes	DISEASE: Hereditary folate malabsorption (HFM) [MIM:229050]: Rare autosomal recessive disorder characterized by impaired intestinal folate absorption with folate deficiency resulting in anemia, hypoimmunoglobulinemia with recurrent infections, and recurrent or chronic diarrhea. In many patients, neurological abnormalities such as seizures or mental retardation become apparent during early childhood, attributed to impaired transport of folates into the central nervous system. When diagnosed early, the disorder can be treated by administration of folate. If untreated, it can be fatal and, if treatment is delayed, the neurological defects can become permanent. Note=The disease is caused by mutations affecting the gene represented in this entry.	cellular iron ion homeostasis [GO:0006879]; folic acid import across plasma membrane [GO:1904447]; folic acid metabolic process [GO:0046655]; folic acid transport [GO:0015884]; hydrogen ion transmembrane transport [GO:1902600]; intestinal folate absorption [GO:0098829]; methotrexate transport [GO:0051958]

152	O15534	PER1_HUMAN	PER1 KIAA0482 PER RIGUI	Period circadian protein homolog 1 (hPER1) (Circadian clock protein PERIOD 1) (Circadian pacemaker protein Rigui)	H409	heme b	Regulatory		Cytoplasm, Nucleus	No		circadian regulation of gene expression [GO:0032922]; circadian regulation of translation [GO:0097167]; circadian rhythm [GO:0007623]; entrainment of circadian clock [GO:0009649]; entrainment of circadian clock by photoperiod [GO:0043153]; histone H3 acetylation [GO:0043966]; histone H3 deacetylation [GO:0070932]; histone H4 acetylation [GO:0043967]; negative regulation of glucocorticoid receptor signaling pathway [GO:2000323]; negative regulation of I-kappaB kinase/NF-kappaB signaling [GO:0043124]; negative regulation of JNK cascade [GO:0046329]; negative regulation of transcription, DNA-templated [GO:0045892]; negative regulation of transcription from RNA polymerase II promoter [GO:0000122]; positive regulation of transcription from RNA polymerase II promoter [GO:0045944]; posttranscriptional regulation of gene expression [GO:0010608]; regulation of circadian rhythm [GO:0042752]; regulation of cytokine production involved in inflammatory response [GO:1900015]; regulation of hair cycle [GO:0042634]; regulation of p38MAPK cascade [GO:1900744]; regulation of sodium ion transport [GO:0002028]; response to cAMP [GO:0051591]; transcription, DNA-templated [GO:0006351]
153	P56645	PER3_HUMAN	PER3 GIG13	Period circadian protein homolog 3 (hPER3) (Cell growth-inhibiting gene 13 protein) (Circadian clock protein PERIOD 3)	Unknown	heme b	Unknown		Cytoplasm, Nucleus	No	DISEASE: Advanced sleep phase syndrome, familial, 3 (FASPS3) [MIM:616882]: A disorder characterized by very early sleep onset and offset. Individuals are 'morning larks' with a 4 hours advance of the sleep, temperature and melatonin rhythms. {ECO:0000269|PubMed:26903630}. Note=The disease is caused by mutations affecting the gene represented in this entry.	circadian regulation of gene expression [GO:0032922]; negative regulation of transcription from RNA polymerase II promoter [GO:0000122]; protein stabilization [GO:0050821]; regulation of circadian sleep/wake cycle, sleep [GO:0045187]; transcription, DNA-templated [GO:0006351]
154	P11678	PERE_HUMAN	EPX EPER EPO EPP	Eosinophil peroxidase (EPO) (EC 1.11.1.7) [Cleaved into: Eosinophil peroxidase light chain; Eosinophil peroxidase heavy chain]	H474	heme i	Catalytic	1.11.1.7	Cytoplasm	No	DISEASE: Eosinophil peroxidase deficiency (EPXD) [MIM:261500]: A rare abnormality without clinical symptoms characterized by decreased or absent peroxidase activity and decreased volume of the granule matrix in eosinophils. {ECO:0000269|PubMed:7809065}. Note=The disease is caused by mutations affecting the gene represented in this entry.	defense response to nematode [GO:0002215]; eosinophil migration [GO:0072677]; hydrogen peroxide catabolic process [GO:0042744]; negative regulation of interleukin-10 production [GO:0032693]; negative regulation of interleukin-5 production [GO:0032714]; neutrophil degranulation [GO:0043312]; positive regulation of interleukin-4 production [GO:0032753]; response to oxidative stress [GO:0006979]
155	P22079	PERL_HUMAN	LPO SAPX	Lactoperoxidase (LPO) (EC 1.11.1.7) (Salivary peroxidase) (SPO)	H468	heme i	Catalytic	1.11.1.7	Extracellular space	No		defense response to bacterium [GO:0042742]; detection of chemical stimulus involved in sensory perception of bitter taste [GO:0001580]; hydrogen peroxide catabolic process [GO:0042744]; response to oxidative stress [GO:0006979]
156	P05164	PERM_HUMAN	MPO	Myeloperoxidase (MPO) (EC 1.11.2.2) [Cleaved into: Myeloperoxidase; 89 kDa myeloperoxidase; 84 kDa myeloperoxidase; Myeloperoxidase light chain; Myeloperoxidase heavy chain]	H502	heme m	Catalytic	1.11.2.2	Unknown	No	DISEASE: Myeloperoxidase deficiency (MPOD) [MIM:254600]: A disorder characterized by decreased myeloperoxidase activity in neutrophils and monocytes that results in disseminated candidiasis. {ECO:0000269|PubMed:7904599, ECO:0000269|PubMed:8142659, ECO:0000269|PubMed:8621627, ECO:0000269|PubMed:9354683, ECO:0000269|PubMed:9637725}. Note=The disease is caused by mutations affecting the gene represented in this entry.	aging [GO:0007568]; defense response [GO:0006952]; defense response to bacterium [GO:0042742]; defense response to fungus [GO:0050832]; hydrogen peroxide catabolic process [GO:0042744]; hypochlorous acid biosynthetic process [GO:0002149]; low-density lipoprotein particle remodeling [GO:0034374]; negative regulation of apoptotic process [GO:0043066]; negative regulation of growth of symbiont in host [GO:0044130]; neutrophil degranulation [GO:0043312]; oxidation-reduction process [GO:0055114]; removal of superoxide radicals [GO:0019430]; respiratory burst involved in defense response [GO:0002679]; response to food [GO:0032094]; response to gold nanoparticle [GO:1990268]; response to lipopolysaccharide [GO:0032496]; response to mechanical stimulus [GO:0009612]; response to oxidative stress [GO:0006979]; response to yeast [GO:0001878]
157	P07202	PERT_HUMAN	TPO	Thyroid peroxidase (TPO) (EC 1.11.1.8)	H494	heme i	Catalytic	1.11.1.8	Unknown	Yes	DISEASE: Note=An alternative splicing in the thyroperoxidase mRNA can cause Graves' disease.; DISEASE: Thyroid dyshormonogenesis 2A (TDH2A) [MIM:274500]: A disorder due to defective conversion of accumulated iodide to organically bound iodine. The iodide organification defect can be partial or complete.Note=The disease is caused by mutations affecting the gene represented in this entry.	embryonic hemopoiesis [GO:0035162]; hormone biosynthetic process [GO:0042446]; hydrogen peroxide catabolic process [GO:0042744]; response to oxidative stress [GO:0006979]; thyroid hormone generation [GO:0006590]

158	Q9H7Z7	PGES2_HUMAN	PTGES2 C9orf15 PGES2	Prostaglandin E synthase 2 (Membrane-associated prostaglandin E synthase-2) (mPGE synthase-2) (Microsomal prostaglandin E synthase 2) (mPGES-2) (Prostaglandin-H(2) E-isomerase) (EC 5.3.99.3) [Cleaved into: Prostaglandin E synthase 2 truncated form]	Unknown	heme b	Unknown	5.3.99.3	Golgi apparatus	Yes		cell redox homeostasis [GO:0045454]; cyclooxygenase pathway [GO:0019371]; neutrophil degranulation [GO:0043312]; positive regulation of transcription, DNA-templated [GO:0045893]
159	P23219	PGH1_HUMAN	PTGS1 COX1	Prostaglandin G/H synthase 1 (EC 1.14.99.1) (Cyclooxygenase-1) (COX-1) (Prostaglandin H2 synthase 1) (PGH synthase 1) (PGHS-1) (PHS 1) (Prostaglandin-endoperoxide synthase 1)	H387	heme b	Catalytic	1.14.99.1	Endoplasmic reticulum	Yes		cyclooxygenase pathway [GO:0019371]; inflammatory response [GO:0006954]; lipid metabolic process [GO:0006629]; prostaglandin biosynthetic process [GO:0001516]; regulation of blood pressure [GO:0008217]; regulation of cell proliferation [GO:0042127]; response to oxidative stress [GO:0006979]; xenobiotic metabolic process [GO:0006805]
160	P35354	PGH2_HUMAN	PTGS2 COX2	Prostaglandin G/H synthase 2 (EC 1.14.99.1) (Cyclooxygenase-2) (COX-2) (PHS II) (Prostaglandin H2 synthase 2) (PGH synthase 2) (PGHS-2) (Prostaglandin-endoperoxide synthase 2)	H374	heme b	Catalytic	1.14.99.1	Endoplasmic reticulum	Yes		aging [GO:0007568]; angiogenesis [GO:0001525]; bone mineralization [GO:0030282]; brown fat cell differentiation [GO:0050873]; cellular response to ATP [GO:0071318]; cellular response to fluid shear stress [GO:0071498]; cellular response to heat [GO:0034605]; cellular response to hypoxia [GO:0071456]; cellular response to lead ion [GO:0071284]; cellular response to mechanical stimulus [GO:0071260]; cellular response to non-ionic osmotic stress [GO:0071471]; cellular response to UV [GO:0034644]; inflammatory response [GO:0006954]; movement of cell or subcellular component [GO:0006928]; negative regulation of calcium ion transport [GO:0051926]; negative regulation of cell cycle [GO:0045786]; negative regulation of cell proliferation [GO:0008285]; negative regulation of cysteine-type endopeptidase activity involved in apoptotic process [GO:0043154]; negative regulation of intrinsic apoptotic signaling pathway in response to osmotic stress [GO:1902219]; positive regulation of apoptotic process [GO:0043065]; positive regulation of brown fat cell differentiation [GO:0090336]; response to fructose [GO:0009750]; response to glucocorticoid [GO:0051384]; response to lipopolysaccharide [GO:0032496]; response to lithium ion [GO:0010226]; response to manganese ion [GO:0010042]; response to oxidative stress [GO:0006979]; response to tumor necrosis factor [GO:0034612]; response to vitamin D [GO:0033280]; sensory perception of pain [GO:0019233]
161	O00264	PGRC1_HUMAN	PGRMC1 HPR6.6 PGRMC	Membrane-associated progesterone receptor component 1 (mPR) (Dap1) (IZA)	Y113	heme b	Electron transfer		Endoplasmic reticulum	Yes		neutrophil degranulation [GO:0043312]
162	O15173	PGRC2_HUMAN	PGRMC2 DG6 PMBP	Membrane-associated progesterone receptor component 2 (Progesterone membrane-binding protein) (Steroid receptor protein DG6)	Y143	heme b	Electron transfer		Unknown	Yes		
163	Q16647	PTGIS_HUMAN	PTGIS CYP8 CYP8A1	Prostacyclin synthase (EC 5.3.99.4) (Prostaglandin I2 synthase)	C441	heme b	Catalytic	5.3.99.4	Endoplasmic reticulum	Yes	DISEASE: Essential hypertension (EHT) [MIM:145500]: A condition in which blood pressure is consistently higher than normal with no identifiable cause. {ECO:0000269|PubMed:12372404}. Note=The disease may be caused by mutations affecting the gene represented in this entry.	apoptotic signaling pathway [GO:0097190]; cellular response to hypoxia [GO:0071456]; cellular response to interleukin-1 [GO:0071347]; cellular response to interleukin-6 [GO:0071354]; cyclooxygenase pathway [GO:0019371]; decidualization [GO:0046697]; embryo implantation [GO:0007566]; icosanoid metabolic process [GO:0006690]; NAD biosynthesis via nicotinamide riboside salvage pathway [GO:0034356]; negative regulation of inflammatory response [GO:0050728]; negative regulation of NF-kappaB transcription factor activity [GO:0032088]; negative regulation of nitric oxide biosynthetic process [GO:0045019]; positive regulation of angiogenesis [GO:0045766]; positive regulation of execution phase of apoptosis [GO:1900119]; positive regulation of peroxisome proliferator activated receptor signaling pathway [GO:0035360]; prostaglandin biosynthetic process [GO:0001516]

164	Q92626	PXDN_HUMAN	PXDN KIAA0230 MG50 PRG2 VPO VPO1	Peroxidasin homolog (EC 1.11.1.7) (Melanoma-associated antigen MG50) (Vascular peroxidase 1) (p53-responsive gene 2 protein)	H1074	heme b	Catalytic	1.11.1.7	Extracellular space, Extracellular space	No	DISEASE: Anterior segment dysgenesis 7 (ASGD7) [MIM:269400]: A form of anterior segment dysgenesis, a group of defects affecting anterior structures of the eye including cornea, iris, lens, trabecular meshwork, and Schlemm canal. Anterior segment dysgeneses result from abnormal migration or differentiation of the neural crest derived mesenchymal cells that give rise to components of the anterior chamber during eye development. Different anterior segment anomalies may exist alone or in combination, including iris hypoplasia, enlarged or reduced corneal diameter, corneal vascularization and opacity, posterior embryotoxon, corectopia, polycoria, abnormal iridocorneal angle, ectopia lentis, and anterior synechiae between the iris and posterior corneal surface. Clinical conditions falling within the phenotypic spectrum of anterior segment dysgeneses include aniridia, Axenfeld anomaly, Reiger anomaly/syndrome, Peters anomaly, and iridogoniodysgenesis. ASGD7 is an autosomal recessive disease. {ECO:0000269|PubMed:21907015}. Note=The disease is caused by mutations affecting the gene represented in this entry.	extracellular matrix organization [GO:0030198]; hydrogen peroxide catabolic process [GO:0042744]; immune response [GO:0006955]; oxidation-reduction process [GO:0055114]; response to oxidative stress [GO:0006979]
165	A1KZ92	PXDNL_HUMAN	PXDNL VPO2	Peroxidasin-like protein (EC 1.11.1.7) (Cardiac peroxidase) (Vascular peroxidase 2) (polysomal ribonuclease 1) (PRM1)	H1057	heme b	Catalytic	1.11.1.7	Extracellular space	No		hydrogen peroxide catabolic process [GO:0042744]; oxidation-reduction process [GO:0055114]; response to oxidative stress [GO:0006979]
166	Q13120	Q13120_HUMAN	CYP2A6V2	Cytochrome P450	Unknown	heme b	Catalytic		Unknown	No		
167	Q14097	Q14097_HUMAN	CYP2B CYP2B7	CYP2B protein (Cytochrome P450 2B7 short isoform)	Unknown	heme b	Catalytic		Unknown	No		
168	Q14412	Q14412_HUMAN	G-gamma HBG1	A-gamma globin (G-gamma globin) (Fragment)	Unknown	heme b	Oxygen storage/transport		Unknown	No		
169	Q16750	Q16750_HUMAN	CYP2C	Unspecific monooxygenase (EC 1.14.14.1) (Fragment)	Unknown	heme b	Catalytic	1.14.14.1	Unknown	No		
170	Q5HYD9	Q5HYD9_HUMAN	DKFZp686M0619	Uncharacterized protein DKFZp686M0619 (Fragment)	Unknown	heme b	Electron transfer		Unknown	No		
171	Q658T6	Q658T6_HUMAN	DKFZp666P073	Uncharacterized protein DKFZp666P073	Unknown	heme b	Catalytic		Unknown	No		
172	Q68D05	Q68D05_HUMAN	DKFZp686G0638	Uncharacterized protein DKFZp686G0638	Unknown	heme b	Catalytic		Unknown	No		
173	Q68D50	Q68D50_HUMAN	DKFZp779I1858	Cytochrome c heme lyase (EC 4.4.1.17)	Unknown	heme c	Substrate - Protein biosynthesis	4.4.1.17	Mitochondrion	Yes		
174	Q6LEN0	Q6LEN0_HUMAN	PGIS	Prostacyclin synthase (EC 5.3.99.4) (Fragment)	Unknown	heme b	Catalytic	5.3.99.4	Unknown	No		prostaglandin biosynthetic process [GO:0001516]
175	Q6ZNJ6	Q6ZNJ6_HUMAN	FLJ00329	FLJ00329 protein (Fragment)	Unknown	heme b	Catalytic		Unknown	No		
176	Q7Z2Y6	Q7Z2Y6_HUMAN	DKFZp686G24255	Uncharacterized protein DKFZp686G24255 (Fragment)	Unknown	heme b	Catalytic		Unknown	No		defense response to bacterium [GO:0042742]; response to oxidative stress [GO:0006979]
177	Q7Z348	Q7Z348_HUMAN	DKFZp686I24235	Uncharacterized protein DKFZp686I24235 (Fragment)	Unknown	heme b	Catalytic		Unknown	No		
178	Q8N3P5	Q8N3P5_HUMAN	DKFZp761K058	Uncharacterized protein DKFZp761K058	Unknown	heme b	Catalytic		Unknown	No		

179	P35398	RORA_HUMAN	RORA NR1F1 RZRA	Nuclear receptor ROR-alpha (Nuclear receptor RZR-alpha) (Nuclear receptor subfamily 1 group F member 1) (RAR-related orphan receptor A) (Retinoid-related orphan receptor-alpha)	H484	heme b	Regulatory - transcription		Nucleus	No		angiogenesis [GO:0001525]; cellular response to hypoxia [GO:0071456]; cellular response to interleukin-1 [GO:0071347]; cellular response to sterol [GO:0036315]; cellular response to tumor necrosis factor [GO:0071356]; cerebellar granule cell precursor proliferation [GO:0021930]; cerebellar Purkinje cell differentiation [GO:0021702]; cGMP metabolic process [GO:0046068]; circadian regulation of gene expression [GO:0032922]; intracellular receptor signaling pathway [GO:0030522]; muscle cell differentiation [GO:0042692]; negative regulation of fat cell differentiation [GO:0045599]; negative regulation of I-kappaB kinase/NF-kappaB signaling [GO:0043124]; negative regulation of inflammatory response [GO:0050728]; nitric oxide biosynthetic process [GO:0006809]; positive regulation of circadian rhythm [GO:0042753]; positive regulation of transcription, DNA-templated [GO:0045893]; positive regulation of transcription from RNA polymerase II promoter [GO:0045944]; positive regulation of vascular endothelial growth factor production [GO:0010575]; regulation of cholesterol homeostasis [GO:2000188]; regulation of glucose metabolic process [GO:0010906]; regulation of macrophage activation [GO:0043030]; regulation of smoothened signaling pathway [GO:0008589]; regulation of steroid metabolic process [GO:0019218]; regulation of transcription, DNA-templated [GO:0006355]; regulation of transcription involved in cell fate commitment [GO:0060850]; T-helper 17 cell differentiation [GO:0072539]; transcription initiation from RNA polymerase II promoter [GO:0006367]; triglyceride homeostasis [GO:0070328]; xenobiotic metabolic process [GO:0006805]
180	Q92753	RORB_HUMAN	RORB NR1F2 RZRB	Nuclear receptor ROR-beta (Nuclear receptor RZR-beta) (Nuclear receptor subfamily 1 group F member 2) (Retinoid-related orphan receptor-beta)	H434	heme b	Regulatory - transcription		Nucleus	No		amacrine cell differentiation [GO:0035881]; cellular response to retinoic acid [GO:0071300]; eye photoreceptor cell development [GO:0042462]; negative regulation of osteoblast differentiation [GO:0045668]; negative regulation of transcription, DNA-templated [GO:0045892]; positive regulation of transcription, DNA-templated [GO:0045893]; regulation of circadian rhythm [GO:0042752]; regulation of transcription, DNA-templated [GO:0006355]; retina development in camera-type eye [GO:0060041]; retinal cone cell development [GO:0046549]; retinal rod cell development [GO:0046548]; rhythmic process [GO:0048511]; transcription initiation from RNA polymerase II promoter [GO:0006367]; visual perception [GO:0007601]
181	P51449	RORG_HUMAN	RORC NR1F3 RORG RZRG	Nuclear receptor ROR-gamma (Nuclear receptor RZR-gamma) (Nuclear receptor subfamily 1 group F member 3) (RAR-related orphan receptor C) (Retinoid-related orphan receptor-gamma)	H479	heme b	Regulatory - transcription		Nucleus	No	DISEASE: Immunodeficiency 42 (IMD42) [MIM:616622]: An autosomal recessive primary immunodeficiency characterized by increased susceptibility to concomitant candidiasis and mycobacteriosis. Candidiasis is characterized by persistent and/or recurrent infections of the skin, nails and mucous membranes caused by organisms of the genus Candida. Mycobacteriosis is characterized by infections caused by moderately virulent mycobacterial species, such as Bacillus Calmette-Guerin (BCG) vaccine, environmental non-tuberculous mycobacteria, and by the more virulent Mycobacterium tuberculosis. IMD42 patients vaccinated with BCG are particularly at risk for developing disseminated mycobacterial infections. {ECO:0000269|PubMed:26160376}. Note=The disease is caused by mutations affecting the gene represented in this entry.	adipose tissue development [GO:0060612]; cellular response to sterol [GO:0036315]; circadian regulation of gene expression [GO:0032922]; lymph node development [GO:0048535]; negative regulation of thymocyte apoptotic process [GO:0070244]; negative regulation of transcription from RNA polymerase II promoter [GO:0000122]; Peyer's patch development [GO:0048541]; positive regulation of circadian rhythm [GO:0042753]; positive regulation of transcription, DNA-templated [GO:0045893]; regulation of fat cell differentiation [GO:0045598]; regulation of glucose metabolic process [GO:0010906]; regulation of steroid metabolic process [GO:0019218]; regulation of transcription involved in cell fate commitment [GO:0060850]; T-helper 17 cell differentiation [GO:0072539]; T-helper cell differentiation [GO:0042093]; transcription initiation from RNA polymerase II promoter [GO:0006367]; xenobiotic metabolic process [GO:0006805]

182	O15357	SHIP2_HUMAN	INPPL1 SHIP2	Phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase 2 (EC 3.1.3.86) (Inositol polyphosphate phosphatase-like protein 1) (INPPL-1) (Protein 51C) (SH2 domain-containing inositol 5'-phosphatase 2) (SH2 domain-containing inositol phosphatase 2) (SHIP-2)	C405	Unknown	Unknown	3.1.3.86	Cytoplasm, Cell membrane	Yes	DISEASE: Diabetes mellitus, non-insulin-dependent (NIDDM) [MIM:125853]: A multifactorial disorder of glucose homeostasis caused by a lack of sensitivity to the body's own insulin. Affected individuals usually have an obese body habitus and manifestations of a metabolic syndrome characterized by diabetes, insulin resistance, hypertension and hypertriglyceridemia. The disease results in long-term complications that affect the eyes, kidneys, nerves, and blood vessels. {ECO:0000269|PubMed:12086927, ECO:0000269|PubMed:15687335}. Note=Disease susceptibility may be associated with variations affecting the gene represented in this entry.; DISEASE: Note=Genetic variations in INPPL1 may be a cause of susceptibility to metabolic syndrome. Metabolic syndrome is characterized by diabetes, insulin resistance, hypertension, and hypertriglyceridemia is absent. {ECO:0000269|PubMed:15220217, ECO:0000269|PubMed:17557929}.; DISEASE: Opsismodysplasia (OPSMD) [MIM:258480]: A rare skeletal dysplasia involving delayed bone maturation. Clinical signs observed at birth include short limbs, small hands and feet, relative macrocephaly with a large anterior fontanel, and characteristic craniofacial abnormalities including a prominent brow, depressed nasal bridge, a small anteverted nose, and a relatively long philtrum. Death secondary to respiratory failure during the first few years of life has been reported, but there can be long-term survival. Typical radiographic findings include shortened long bones with very delayed epiphyseal ossification, severe platyspondyly, metaphyseal cupping, and characteristic abnormalities of the metacarpals and phalanges. {ECO:0000269|PubMed:23273569}. Note=The disease is caused by mutations affecting the gene represented in this entry.	actin filament organization [GO:0007015]; cell adhesion [GO:0007155]; endochondral ossification [GO:0001958]; endocytosis [GO:0006897]; glucose metabolic process [GO:0006006]; immune system process [GO:0002376]; inositol phosphate metabolic process [GO:0043647]; negative regulation of cell proliferation [GO:0008285]; negative regulation of gene expression [GO:0010629]; phosphatidylinositol biosynthetic process [GO:0006661]; phosphatidylinositol dephosphorylation [GO:0046856]; post-embryonic development [GO:0009791]; response to insulin [GO:0032868]; ruffle assembly [GO:0097178]
183	P12931	SRC_HUMAN	SRC SRC1	Proto-oncogene tyrosine-protein kinase Src (EC 2.7.10.2) (Proto-oncogene c-Src) (pp60c-src) (p60-Src)	Unknown	Unknown	Regulatory	2.7.10.2	Cytoplasm, Mitochondrion, Cell membrane, Nucleus	Yes	DISEASE: Note=SRC kinase activity has been shown to be increased in several tumor tissues and tumor cell lines such as colon carcinoma cells. {ECO:0000269|PubMed:2498394, ECO:0000269|PubMed:3093483}.; DISEASE: Thrombocytopenia 6 (THC6) [MIM:616937]: A form of thrombocytopenia, a hematologic disorder defined by a decrease in the number of platelets in circulating blood, resulting in the potential for increased bleeding and decreased ability for clotting. THC6 is an autosomal dominant form. Affected individuals may also have bone abnormalities and an increased risk for myelofibrosis. {ECO:0000269|PubMed:26936507}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cell cycle [GO:0007049]; cell proliferation [GO:0008283]; cellular response to fatty acid [GO:0071398]; cellular response to fluid shear stress [GO:0071498]; cellular response to hypoxia [GO:0071456]; cellular response to insulin stimulus [GO:0032869]; cellular response to lipopolysaccharide [GO:0071222]; cellular response to reactive oxygen species [GO:0034614]; central nervous system development [GO:0007417]; negative regulation of apoptotic process [GO:0043066]; negative regulation of cysteine-type endopeptidase activity involved in apoptotic process [GO:0043154]; negative regulation of extrinsic apoptotic signaling pathway [GO:2001237]; negative regulation of focal adhesion assembly [GO:0051895]; negative regulation of intrinsic apoptotic signaling pathway [GO:2001243]; negative regulation of mitochondrial depolarization [GO:0051902]; positive regulation of apoptotic process [GO:0043065]; positive regulation of cytokine secretion [GO:0050715]; positive regulation of integrin activation [GO:0033625]; positive regulation of lamellipodium morphogenesis [GO:2000394]; positive regulation of protein localization to nucleus [GO:1900182]; positive regulation of protein processing [GO:0010954]; regulation of cell proliferation [GO:0042127]; regulation of early endosome to late endosome transport [GO:2000641]; regulation of podosome assembly [GO:0071801]; regulation of protein binding [GO:0043393]; regulation of vascular permeability [GO:0043114]; response to acidic pH [GO:0010447]; response to drug [GO:0042493]; response to electrical stimulus [GO:0051602]; response to hydrogen peroxide [GO:0042542];

184	O76061	STC2_HUMAN	STC2	Stanniocalcin-2 (STC-2) (Stanniocalcin-related protein) (STC-related protein) (STCRP)	Unknown	heme b	Substrate - Regulatory/Sensor		Extracellular space	No		cellular calcium ion homeostasis [GO:0006874]; cellular protein metabolic process [GO:0044267]; cellular response to hypoxia [GO:0071456]; decidualization [GO:0046697]; embryo implantation [GO:0007566]; endoplasmic reticulum unfolded protein response [GO:0030968]; negative regulation of gene expression [GO:0010629]; negative regulation of multicellular organism growth [GO:0040015]; post-translational protein modification [GO:0043687]; regulation of hormone biosynthetic process [GO:0046885]; regulation of store-operated calcium entry [GO:2001256]; response to oxidative stress [GO:0006979]; response to peptide hormone [GO:0043434]; response to vitamin D [GO:0033280]
185	Q9UHE8	STEA1_HUMAN	STEAP1 PRSS24 STEAP	Metalloreductase STEAP1 (EC 1.16.1.-) (Six-transmembrane epithelial antigen of prostate 1)	H175-H268	heme b	Electron transfer	1.16.1.-	Endosome	Yes		ion transport [GO:0006811]; iron ion homeostasis [GO:0055072]
186	Q8NFT2	STEA2_HUMAN	STEAP2 PCANAP1 STAMP1 UNQ6507/PRO23203	Metalloreductase STEAP2 (EC 1.16.1.-) (Prostate cancer-associated protein 1) (Protein up-regulated in metastatic prostate cancer) (PUMPCn) (Six-transmembrane epithelial antigen of prostate 2) (SixTransMembrane protein of prostate 1)	H316-H409	heme b	Electron transfer	1.16.1.-	Cell membrane, Endosome	Yes		copper ion import [GO:0015677]; endocytosis [GO:0006897]; ferric iron import across plasma membrane [GO:0098706]; Golgi to plasma membrane transport [GO:0006893]; iron ion homeostasis [GO:0055072]; regulated exocytosis [GO:0045055]; response to hormone [GO:0009725]
187	Q658P3	STEA3_HUMAN	STEAP3 TSAP6	Metalloreductase STEAP3 (EC 1.16.1.-) (Dudulin-2) (Six-transmembrane epithelial antigen of prostate 3) (Tumor suppressor-activated pathway protein 6) (hTSAP6) (pHyde) (hpHyde)	H316-H409	heme b	Electron transfer	1.16.1.-	Nucleus, Endosome, Cell membrane	Yes	DISEASE: Anemia, hypochromic microcytic, with iron overload 2 (AHMIO2) [MIM:615234]: A hematologic disease characterized by abnormal hemoglobin content in the erythrocytes which are reduced in size, severe anemia, erythropoietic hyperplasia of bone marrow, massive hepatic iron deposition, and hepatosplenomegaly. {ECO:0000269|PubMed:22031863}. Note=The disease is caused by mutations affecting the gene represented in this entry.	apoptotic process [GO:0006915]; cell cycle [GO:0007049]; copper ion import [GO:0015677]; ferric iron import across plasma membrane [GO:0098706]; iron ion homeostasis [GO:0055072]; protein secretion [GO:0009306]; regulation of apoptotic process [GO:0042981]; transferrin transport [GO:0033572]
188	Q687X5	STEA4_HUMAN	STEAP4 STAMP2 TNFAIP9	Metalloreductase STEAP4 (EC 1.16.1.-) (Six-transmembrane epithelial antigen of prostate 4) (SixTransMembrane protein of prostate 2) (Tumor necrosis factor, alpha-induced protein 9)	H304-H397	heme b	Electron transfer	1.16.1.-	Golgi apparatus, Cell membrane, Endosome	Yes		copper ion import [GO:0015677]; fat cell differentiation [GO:0045444]; ferric iron import across plasma membrane [GO:0098706]; iron ion homeostasis [GO:0055072]
189	P51687	SUOX_HUMAN	SUOX	Sulfite oxidase, mitochondrial (EC 1.8.3.1)	H118-H143	heme b	Electron transfer	1.8.3.1	Mitochondrion	Yes	DISEASE: Isolated sulfite oxidase deficiency (ISOD) [MIM:272300]: Characterized by neurological abnormalities including multicystic leukoencephalopathy with brain atrophy. Patients often suffer from seizures. Often leads to death at an early age. Note=The disease is caused by mutations affecting the gene represented in this entry.	nitrate assimilation [GO:0042128]; sulfide oxidation, using sulfide:quinone oxidoreductase [GO:0070221]
190	P48775	T23O_HUMAN	TDO2 TDO	Tryptophan 2,3-dioxygenase (TDO) (EC 1.13.11.11) (Tryptamin 2,3-dioxygenase) (Tryptophan oxygenase) (TO) (TRPO) (Tryptophan pyrrolase) (Tryptophanase)	H328	heme b	Catalytic	1.13.11.11	Unknown	No		protein homotetramerization [GO:0051289]; tryptophan catabolic process [GO:0006569]; tryptophan catabolic process to acetyl-CoA [GO:0019442]; tryptophan catabolic process to kynurenine [GO:0019441]
191	Q8WY91	THAP4_HUMAN	THAP4 CGI-36 PP238	THAP domain-containing protein 4	H567	heme b	Regulatory - transcription		Unknown	No		
192	P24557	THAS_HUMAN	TBXAS1 CYP5 CYP5A1	Thromboxane-A synthase (TXA synthase) (TXS) (EC 5.3.99.5) (Cytochrome P450 5A1)	C479	heme b	Catalytic	5.3.99.5	Endoplasmic reticulum	Yes	DISEASE: Ghosal hematodiaphyseal dysplasia (GHDD) [MIM:231095]: Rare autosomal recessive disorder characterized by increased bone density with predominant diaphyseal involvement and aregenerative corticosteroid-sensitive anemia. Aregenerative anemia is characterized by bone marrow failure, so that functional marrow cells are regenerated slowly or not at all. {ECO:0000269|PubMed:18264100}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Note=Thromboxane synthetase deficiency has been detected in some patients with a bleeding disorder due to platelet dysfunction. {ECO:0000269|PubMed:6101498}.	cyclooxygenase pathway [GO:0019371]; icosanoid metabolic process [GO:0006690]

Table S6: *Functional properties of the human iron-sulfur proteins.*

	Uniprot Id	Entry name	Gene names	Protein names	Predicted Pattern	Number of cofactors	Iron-cofactor role	EC number	Subcellular location	Membrane associated	Involvement in disease	Gene ontology (biological process)
1	O75027	ABCB7_HUMAN	ABCB7 ABC7	ATP-binding cassette sub-family B member 7, mitochondrial (ATP-binding cassette transporter 7) (ABC transporter 7 protein)	Unknown	Fe_2S_2	Substrate - transport		Mitochondrion	Yes	DISEASE: Anemia, sideroblastic, spinocerebellar ataxia (ASAT) [MIM:301310]: A X-linked recessive disorder characterized by an infantile to early childhood onset of non-progressive cerebellar ataxia and mild anemia, with hypochromia and microcytosis. {ECO:0000269|PubMed:10196363, ECO:0000269|PubMed:11050011, ECO:0000269|PubMed:11843825, ECO:0000269|PubMed:22398176}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cellular iron ion homeostasis [GO:0006879]; transmembrane transport [GO:0055085]; transport [GO:0006810]
2	P61221	ABCE1_HUMAN	ABCE1 RLI RNASEL1 RNASELI RNS4I OK/SW-cl.40	ATP-binding cassette sub-family E member 1 (2'-5'-oligoadenylate-binding protein) (HuHP68) (RNase L inhibitor) (Ribonuclease 4 inhibitor) (RNS4I)	C16-C21-C25-C29-C55-C58-C61-C65	2 × Fe_4S_4	Unknown		Cytoplasm, Mitochondrion, Cell membrane	Yes		negative regulation of endoribonuclease activity [GO:0060702]; regulation of type I interferon-mediated signaling pathway [GO:0060338]; ribosomal subunit export from nucleus [GO:0000054]; translational initiation [GO:0006413]; translational termination [GO:0006415]; viral process [GO:0016032]
3	P21399	ACOC_HUMAN	ACO1 IREB1	Cytoplasmic aconitate hydratase (Aconitase) (EC 4.2.1.3) (Citrate hydro-lyase) (Ferritin repressor protein) (Iron regulatory protein 1) (IRP1) (Iron-responsive element-binding protein 1) (IRE-BP 1)	C437-C503-C506	Fe_4S_4	Substrate - sensor	4.2.1.3	Cytoplasm	No		cellular iron ion homeostasis [GO:0006879]; citrate metabolic process [GO:0006101]; intestinal absorption [GO:0050892]; post-embryonic development [GO:0009791]; regulation of translation [GO:0006417]; response to iron(II) ion [GO:0010040]; tricarboxylic acid cycle [GO:0006099]
4	Q99798	ACON_HUMAN	ACO2	Aconitate hydratase, mitochondrial (Aconitase) (EC 4.2.1.3) (Citrate hydro-lyase)	C385-C448-C451	Fe_4S_4	Unknown	4.2.1.3	Mitochondrion	No	DISEASE: Infantile cerebellar-retinal degeneration (ICRD) [MIM:614559]: A severe autosomal recessive neurodegenerative disorder characterized by onset between ages 2 and 6 months of truncal hypotonia, athetosis, seizures, and ophthalmologic abnormalities, particularly optic atrophy and retinal degeneration. Affected individuals show profound psychomotor retardation, with only some achieving rolling, sitting, or recognition of family. Brain MRI shows progressive cerebral and cerebellar degeneration. {ECO:0000269|PubMed:22405087, ECO:0000269|PubMed:25351951}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Optic atrophy 9 (OPA9) [MIM:616289]: A condition that features progressive visual loss in association with optic atrophy. Atrophy of the optic disk indicates a deficiency in the number of nerve fibers which arise in the retina and converge to form the optic disk, optic nerve, optic chiasm and optic tracts. {ECO:0000269|PubMed:25351951}. Note=The disease is caused by mutations affecting the gene represented in this entry.	citrate metabolic process [GO:0006101]; generation of precursor metabolites and energy [GO:0006091]; isocitrate metabolic process [GO:0006102]; liver development [GO:0001889]; response to isolation stress [GO:0035900]; tricarboxylic acid cycle [GO:0006099]
5	P10109	ADX_HUMAN	FDX1 ADX	Adrenodoxin, mitochondrial (Adrenal ferredoxin) (Ferredoxin-1) (Hepatoredoxin)	C106-C112-C115-C152	Fe_2S_2	Electron transfer		Mitochondrion	No		C21-steroid hormone biosynthetic process [GO:0006700]; cellular response to cAMP [GO:0071320]; cellular response to forskolin [GO:1904322]; cholesterol metabolic process [GO:0008203]; hormone biosynthetic process [GO:0042446]; small molecule metabolic process [GO:0044281]; sterol metabolic process [GO:0016125]
6	Q96NN9	AIFM3_HUMAN	AIFM3 AIFL	Apoptosis-inducing factor 3 (EC 1.-.-.-) (Apoptosis-inducing factor-like protein)	C109-H111-C128-H131	Fe_2S_2 (predicted)	Unknown	1.-.-.-	Mitochondrion, Nucleus	No		execution phase of apoptosis [GO:0097194]
7	Q06278	AOXA_HUMAN	AOX1 AO	Aldehyde oxidase (EC 1.2.3.1) (Aldehyde oxidase 1) (Azaheterocycle hydroxylase) (EC 1.17.3.-)	C44-C49-C52-C74; C114-C117-C149-C151	2 × Fe_2S_2	Electron transfer	1.2.3.1; 1.17.3.-	Cytoplasm	No		drug metabolic process [GO:0017144]; oxidation-reduction process [GO:0055114]; vitamin B6 metabolic process [GO:0042816]; xanthine catabolic process [GO:0009115]
8	Q9Y3E2	BOLA1_HUMAN	BOLA1 CGI-143	BolA-like protein 1 (hBolA)	Unknown	Fe_2S_2 shared with GLRX	Substrate - biosinthesis		Mitochondrion	No		
9	Q9H3K6	BOLA2_HUMAN	BOLA2 BOLA2A My016; BOLA2B	BolA-like protein 2	Unknown	Fe_2S_2 shared with GLRX	Substrate - biosinthesis		Cytoplasm, Nucleus	No		[2Fe-2S] cluster assembly [GO:0044571]; interleukin-12-mediated signaling pathway [GO:0035722]; protein maturation by iron-sulfur cluster transfer [GO:0097428]
10	Q53S33	BOLA3_HUMAN	BOLA3	BolA-like protein 3	Unknown	Fe_2S_2 shared with GLRX	Substrate - biosinthesis		Mitochondrion	No	DISEASE: Multiple mitochondrial dysfunctions syndrome 2 with hyperglycinemia (MMDS2) [MIM:614299]: A severe disorder of systemic energy metabolism, resulting in weakness, respiratory failure, lack of neurologic development, lactic acidosis, hyperglycinemia and early death. Some patients show failure to thrive, pulmonary hypertension, hypotonia and irritability. Biochemical features include severe combined deficiency of the 2-oxoacid dehydrogenases, defective lipoic acid synthesis and reduction in activity of mitochondrial respiratory chain complexes. {ECO:0000269|PubMed:21944046, ECO:0000269|PubMed:22562699, ECO:0000269|PubMed:24334290, ECO:0000269|PubMed:26741492}. Note=The disease is caused by mutations affecting the gene represented in this entry.	

11	Q5VV42	CDKAL_HUMAN	CDKAL1	Threonylcarbamoyladenosine tRNA methylthiotransferase (EC 2.8.4.5) (CDK5 regulatory subunit-associated protein 1-like 1) (tRNA-t(6)A37 methylthiotransferase)	C73-C109-C138; C214-C218-C221	2 × Fe_4S_4	Catalytic	2.8.4.5	Endoplasmic reticulum	Yes	DISEASE: Diabetes mellitus, non-insulin-dependent (NIDDM) [MIM:125853]: A multifactorial disorder of glucose homeostasis caused by a lack of sensitivity to the body's own insulin. Affected individuals usually have an obese body habitus and manifestations of a metabolic syndrome characterized by diabetes, insulin resistance, hypertension and hypertriglyceridemia. The disease results in long-term complications that affect the eyes, kidneys, nerves, and blood vessels. {ECO:0000269|PubMed:17460697, ECO:0000269|PubMed:17463246}. Note=Disease susceptibility is associated with variations affecting the gene represented in this entry.	maintenance of translational fidelity [GO:1990145]; tRNA modification [GO:0006400]
12	Q9NZ45	CISD1_HUMAN	CISD1 C10orf70 ZCD1 MDS029	CDGSH iron-sulfur domain-containing protein 1 (MitoNEET)	C72-C74-C83-H87	Fe_2S_2	Substrate - biogenesis		Mitochondrion	Yes		regulation of cellular respiration [GO:0043457]
13	Q8N5K1	CISD2_HUMAN	CISD2 CDGSH2 ERIS ZCD2	CDGSH iron-sulfur domain-containing protein 2 (Endoplasmic reticulum intermembrane small protein) (MitoNEET-related 1 protein) (Miner1) (Nutrient-deprivation autophagy factor-1) (NAF-1)	C99-C101-C110-H114	Fe_2S_2	Unknown		Endoplasmic reticulum, Mitochondrion	Yes	DISEASE: Wolfram syndrome 2 (WFS2) [MIM:604928]: A rare disorder characterized by juvenile-onset insulin-dependent diabetes mellitus with optic atrophy. Other manifestations include diabetes insipidus, sensorineural deafness, dementia, psychiatric illnesses. WFS2 patients additionally show a strong bleeding tendency and gastrointestinal ulceration. Diabetes insipidus may be absent. {ECO:0000269|PubMed:17846994}. Note=The disease is caused by mutations affecting the gene represented in this entry.	autophagy of mitochondrion [GO:0000422]; multicellular organism aging [GO:0010259]; regulation of autophagy [GO:0010506]
14	P0C7P0	CISD3_HUMAN	CISD3	CDGSH iron-sulfur domain-containing protein 3, mitochondrial (MitoNEET-related protein 2) (Miner2)	C60-C62-C71-H75; C98-C100-C109-H113	2 × Fe_2S_2	Unknown		Mitochondrion	No		
15	Q96SZ6	CK5P1_HUMAN	CDK5RAP1 C20orf34 CGI-05 HSPC167	CDK5 regulatory subunit-associated protein 1 (CDK5 activator-binding protein C42)	C109-C145-C183; C258-C262-C265	2 × Fe_4S_4	Catalytic		Unknown	No		brain development [GO:0007420]; mitochondrial tRNA modification [GO:0070900]; negative regulation of cyclin-dependent protein serine/threonine kinase activity [GO:0045736]; positive regulation of mitochondrial translation [GO:0070131]; positive regulation of translational fidelity [GO:0045903]; regulation of neuron differentiation [GO:0045664]
16	Q9Y471	CMAH_HUMAN	CMAHP CMAH	Inactive cytidine monophosphate-N-acetylneuraminic acid hydroxylase (CMP-NeuAc hydroxylase-like protein) (Cytidine monophosphate-N-acetylneuraminic acid hydroxylase pseudogene)	Unknown	Fe_2S_2	Unknown		Cytoplasm, Nucleus	Yes		regulation of Wnt signaling pathway [GO:0030111]
17	Q6FI81	CPIN1_HUMAN	CIAPIN1 CUA001 PRO0915	Anamorsin (Cytokine-induced apoptosis inhibitor 1) (Fe-S cluster assembly protein DRE2 homolog)	C237-C246-C249-C251	2 × Fe_2S_2	Substrate - biogenesis		Cytoplasm, Mitochondrion, Nucleus	Yes		apoptotic process [GO:0006915]; hemopoiesis [GO:0030097]; iron-sulfur cluster assembly [GO:0016226]; negative regulation of apoptotic process [GO:0043066]
18	Q96FC9	DDX11_HUMAN	DDX11 CHL1 CHLR1 KRG2	ATP-dependent DNA helicase DDX11 (EC 3.6.4.12) (CHL1-related protein 1) (hCHLR1) (DEAD/H-box protein 11) (Keratinocyte growth factor-regulated gene 2 protein) (KRG-2)	C267-C285-C315-C350	Fe_4S_4	Structural - Regulatory	3.6.4.12	Cytoplasm, Nucleus	No	DISEASE: Warsaw breakage syndrome (WBRS) [MIM:613398]: A syndrome characterized by severe microcephaly, pre- and postnatal growth retardation, facial dysmorphism and abnormal skin pigmentation. Additional features include high arched palate, coloboma of the right optic disk, deafness, ventricular septal defect, toes and fingers abnormalities. At cellular level, drug-induced chromosomal breakage, a feature of Fanconi anemia, and sister chromatid cohesion defects, a feature of Roberts syndrome, coexist. {ECO:0000269|PubMed:20137776, ECO:0000269|PubMed:23033317, ECO:0000269|PubMed:26089203}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cellular response to bleomycin [GO:1904976]; cellular response to cisplatin [GO:0072719]; cellular response to DNA damage stimulus [GO:0006974]; cellular response to hydroxyurea [GO:0072711]; DNA duplex unwinding [GO:0032508]; DNA repair [GO:0006281]; G-quadruplex DNA unwinding [GO:0044806]; IRE1-mediated unfolded protein response [GO:0036498]; multicellular organism development [GO:0007275]; negative regulation of protein binding [GO:0032091]; nucleolar chromatin organization [GO:1990700]; positive regulation of chromatin binding [GO:0035563]; positive regulation of double-strand break repair [GO:2000781]; positive regulation of endodeoxyribonuclease activity [GO:0032079]; positive regulation of sister chromatid cohesion [GO:0045876]; positive regulation of transcription of nuclear large rRNA transcript from RNA polymerase I promoter [GO:1901838]; replication fork processing [GO:0031297]; sister chromatid cohesion [GO:0007062]; transcription, DNA-templated [GO:0006351]; viral process [GO:0016032]
19	Q92771	DDX12_HUMAN	DDX12P CHLR2 DDX12	Putative ATP-dependent RNA helicase DDX12 (EC 3.6.4.13) (CHL1-related protein 2) (hCHLR2) (DEAD/H box protein 12)	C286-C304-C334-C369	Fe_4S_4	Structural - Regulatory	3.6.4.13	Nucleus	No		cell cycle [GO:0007049]; nucleobase-containing compound metabolic process [GO:0006139]
20	P51530	DNA2_HUMAN	DNA2 DNA2L KIAA0083	DNA replication ATP-dependent helicase/nuclease DNA2 (hDNA2) (DNA replication ATP-dependent helicase-like homolog) [Includes: DNA replication nuclease DNA2 (EC 3.1.-.-); DNA replication ATP-dependent helicase DNA2 (EC 3.6.4.12)]	C136-C393-C396-C402	Fe_4S_4	Structural - Regulatory	3.1.-.-; 3.6.4.12	Mitochondrion, Nucleus	No	DISEASE: Progressive external ophthalmoplegia with mitochondrial DNA deletions, autosomal dominant, 6 (PEOA6) [MIM:615156]: A disorder characterized by muscle weakness, mainly affecting the lower limbs, external ophthalmoplegia, exercise intolerance, and mitochondrial DNA deletions on muscle biopsy. Symptoms may appear in childhood or adulthood and show slow progression. {ECO:0000269|PubMed:23352259}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Seckel syndrome 8 (SCKL8) [MIM:615807]: A rare autosomal recessive disorder characterized by proportionate dwarfism of prenatal onset associated with low birth weight, growth retardation, severe microcephaly with a bird-headed like appearance, and mental retardation. {ECO:0000269|PubMed:24389050}. Note=The disease is caused by mutations affecting the gene represented in this entry.	base-excision repair [GO:0006284]; DNA double-strand break processing [GO:0000729]; DNA replication [GO:0006260]; DNA replication, Okazaki fragment processing [GO:0033567]; DNA replication, removal of RNA primer [GO:0043137]; DNA replication checkpoint [GO:0000076]; DNA synthesis involved in DNA repair [GO:0000731]; G-quadruplex DNA unwinding [GO:0044806]; mitochondrial DNA repair [GO:0043504]; mitochondrial DNA replication [GO:0006264]; mitotic telomere maintenance via semi-conservative replication [GO:1902990]; nucleic acid phosphodiester bond hydrolysis [GO:0090305]; positive regulation of DNA replication [GO:0045740]; regulation of signal transduction by p53 class mediator [GO:1901796]; strand displacement [GO:0000732]; t-circle formation [GO:0090656]; telomere maintenance [GO:0000723]; telomere maintenance via semi-conservative replication [GO:0032201]

21	Q9BZG8	DPH1_HUMAN	DPH1 DPH2L DPH2L1 OVCA1	2-(3-amino-3-carboxypropyl)histidine synthase subunit 1 (EC 2.5.1.108) (Diphthamide biosynthesis protein 1) (Diphtheria toxin resistance protein 1) (Ovarian cancer-associated gene 1 protein) (S-adenosyl-L-methionine:L-histidine 3-amino-3-carboxypropyltransferase 1)	C115-C219-C347	Fe_4S_4	Catalytic	2.5.1.108	Cytoplasm, Nucleus	No	DISEASE: Developmental delay with short stature, dysmorphic features, and sparse hair (DEDSSH) [MIM:616901]: An autosomal recessive syndrome characterized by intellectual disability, short stature, and craniofacial and ectodermal anomalies including scaphocephaly with or without craniosynostosis, prominent forehead, sparse eyebrows and hair, hypoplastic toenails and, in some cases, dental anomalies. {ECO:0000269|PubMed:25558065, ECO:0000269|PubMed:26220823}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cell proliferation [GO:0008283]; peptidyl-diphthamide biosynthetic process from peptidyl-histidine [GO:0017183]
22	Q9BQC3	DPH2_HUMAN	DPH2 DPH2L2	2-(3-amino-3-carboxypropyl)histidine synthase subunit 2 (EC 2.5.1.108) (Diphthamide biosynthesis protein 2) (Diphtheria toxin resistance protein 2) (S-adenosyl-L-methionine:L-histidine 3-amino-3-carboxypropyltransferase 2)	C88-C341	Fe_4S_4	Catalytic	2.5.1.108	Unknown	No		peptidyl-diphthamide biosynthetic process from peptidyl-histidine [GO:0017183]
23	P28340	DPOD1_HUMAN	POLD1 POLD	DNA polymerase delta catalytic subunit (EC 2.7.7.7) (EC 3.1.11.-) (DNA polymerase subunit delta p125)	C1058-C1061-C1071-C1076	Fe_4S_4	Structural - Regulatory	2.7.7.7; 3.1.11.-	Nucleus	No	DISEASE: Colorectal cancer 10 (CRCS10) [MIM:612591]: A complex disease characterized by malignant lesions arising from the inner wall of the large intestine (the colon) and the rectum. Genetic alterations are often associated with progression from premalignant lesion (adenoma) to invasive adenocarcinoma. Risk factors for cancer of the colon and rectum include colon polyps, long-standing ulcerative colitis, and genetic family history. {ECO:0000269|PubMed:23263490, ECO:0000269|PubMed:24501277}. Note=Disease susceptibility is associated with variations affecting the gene represented in this entry.; DISEASE: Mandibular hypoplasia, deafness, progeroid features, and lipodystrophy syndrome (MDPL) [MIM:615381]: An autosomal dominant systemic disorder characterized by prominent loss of subcutaneous fat, metabolic abnormalities including insulin resistance and diabetes mellitus, sclerodermatous skin, and a facial appearance characterized by mandibular hypoplasia. Sensorineural deafness occurs late in the first or second decades of life. {ECO:0000269|PubMed:23770608}. Note=The disease is caused by mutations affecting the gene represented in this entry.	base-excision repair, gap-filling [GO:0006287]; cellular response to UV [GO:0034644]; DNA damage response, detection of DNA damage [GO:0042769]; DNA ligation [GO:0006266]; DNA repair [GO:0006281]; DNA replication [GO:0006260]; DNA replication proofreading [GO:0045004]; DNA synthesis involved in DNA repair [GO:0000731]; fatty acid homeostasis [GO:0055089]; mismatch repair [GO:0006298]; nucleotide-excision repair, DNA gap filling [GO:0006297]; nucleotide-excision repair, DNA incision [GO:0033683]; nucleotide-excision repair, DNA incision, 5'-to lesion [GO:0006296]; response to UV [GO:0009411]; telomere maintenance [GO:0000723]; telomere maintenance via semi-conservative replication [GO:0032201]; transcription-coupled nucleotide-excision repair [GO:0006283]; translesion synthesis [GO:0019985]
24	Q07864	DPOE1_HUMAN	POLE POLE1	DNA polymerase epsilon catalytic subunit A (EC 2.7.7.7) (DNA polymerase II subunit A)	C2221-C2224-C2236-C2238	Fe_4S_4	Structural - Regulatory	2.7.7.7	Nucleus	No	DISEASE: Colorectal cancer 12 (CRCS12) [MIM:615083]: A complex disease characterized by malignant lesions arising from the inner wall of the large intestine (the colon) and the rectum. Genetic alterations are often associated with progression from premalignant lesion (adenoma) to invasive adenocarcinoma. Risk factors for cancer of the colon and rectum include colon polyps, long-standing ulcerative colitis, and genetic family history. CRCS12 is characterized by a high-penetrance predisposition to the development of colorectal adenomas and carcinomas, with a variable tendency to develop multiple and large tumors. Onset is usually before age 40 years. The histologic features of the tumors are unremarkable. {ECO:0000269|PubMed:23263490, ECO:0000269|PubMed:24501277, ECO:0000269|PubMed:25860647, ECO:0000269|PubMed:27573199}. Note=Disease susceptibility is associated with variations affecting the gene represented in this entry.; DISEASE: Facial dysmorphism, immunodeficiency, livedo, and short stature (FILS) [MIM:615139]: A syndrome characterized by mild facial dysmorphism, mainly malar hypoplasia, livedo on the skin since birth, and immunodeficiency resulting in recurrent infections. Growth impairment is observed during early childhood and results in variable short stature in adulthood. {ECO:0000269|PubMed:23230001}. Note=The disease is caused by mutations affecting the gene represented in this entry.	base-excision repair, gap-filling [GO:0006287]; DNA replication [GO:0006260]; DNA replication initiation [GO:0006270]; DNA replication proofreading [GO:0045004]; DNA synthesis involved in DNA repair [GO:0000731]; embryonic organ development [GO:0048568]; G1/S transition of mitotic cell cycle [GO:0000082]; leading strand elongation [GO:0006272]; nucleotide-excision repair, DNA gap filling [GO:0006297]; telomere maintenance via semi-conservative replication [GO:0032201]

25	P09884	DPOLA_HUMAN	POLA1 POLA	DNA polymerase alpha catalytic subunit (EC 2.7.7.7) (DNA polymerase alpha catalytic subunit p180)	C1348-C1353-C1371-C1374	Fe_4S_4	Structural - Regulatory	2.7.7.7	Cytoplasm, Nucleus	No	DISEASE: Pigmentary disorder, reticulate, with systemic manifestations, X-linked (PDR) [MIM:301220]: A X-linked recessive disorder characterized by recurrent infections and sterile inflammation in various organs. Diffuse skin hyperpigmentation with a distinctive reticulate pattern is universally evident by early childhood. This is later followed in many patients by hypohidrosis, corneal inflammation and scarring, enterocolitis that resembles inflammatory bowel disease, and recurrent urethral strictures. Melanin and amyloid deposition is present in the dermis. Affected males also have a characteristic facies with frontally upswept hair and flared eyebrows. Female carriers have only restricted pigmentary changes along Blaschko's lines. {ECO:0000269|PubMed:27019227}. Note=The disease is caused by mutations affecting the gene represented in this entry. XLPDR is caused by a recurrent intronic mutation that results in missplicing and reduced POLA1 expression. This leads to a decrease in cytosolic RNA:DNA hybrids and constitutive activation of type I interferon responses, but has no effect on cell replication. {ECO:0000269|PubMed:27019227}.	cell proliferation [GO:0008283]; DNA replication [GO:0006260]; DNA replication, synthesis of RNA primer [GO:0006269]; DNA replication initiation [GO:0006270]; DNA strand elongation involved in DNA replication [GO:0006271]; double-strand break repair via nonhomologous end joining [GO:0006303]; G1/S transition of mitotic cell cycle [GO:0000082]; lagging strand elongation [GO:0006273]; leading strand elongation [GO:0006272]; regulation of transcription involved in G1/S transition of mitotic cell cycle [GO:0000083]; telomere maintenance via semi-conservative replication [GO:0032201]; viral process [GO:0016032]
26	Q12882	DPYD_HUMAN	DPYD	Dihydropyrimidine dehydrogenase [NADP(+)] (DHPDHase) (DPD) (EC 1.3.1.2) (Dihydrothymine dehydrogenase) (Dihydrouracil dehydrogenase)	C79-C82-C87-C91; C130-C136-C140-C156; C953-C956-C959-C963; C986-C989-C992-C996	4 × Fe_4S_4	Unknown	1.3.1.2	Cytoplasm	No	DISEASE: Dihydropyrimidine dehydrogenase deficiency (DPYDD) [MIM:274270]: A metabolic disorder with large phenotypic variability, ranging from no symptoms to a convulsive disorder with motor and mental retardation. It is characterized by persistent urinary excretion of excessive amounts of uracil, thymine and 5-hydroxymethyluracil. Patients suffering from this disease show a severe reaction to the anticancer drug 5-fluorouracil. {ECO:0000269|PubMed:14702039, ECO:0000269|PubMed:16710414, ECO:0000269|PubMed:9266349, ECO:0000269|PubMed:9439663}. Note=The disease is caused by mutations affecting the gene represented in this entry.	beta-alanine biosynthetic process [GO:0019483]; purine nucleobase catabolic process [GO:0006145]; pyrimidine nucleobase catabolic process [GO:0006208]; pyrimidine nucleoside catabolic process [GO:0046135]; thymidine catabolic process [GO:0006214]; thymine catabolic process [GO:0006210]; uracil catabolic process [GO:0006212]
27	Q9H9T3	ELP3_HUMAN	ELP3	Elongator complex protein 3 (hELP3) (EC 2.3.1.48)	C99-C109-C112	Fe_4S_4	Catalytic	2.3.1.48	Cytoplasm	No	DISEASE: Note=ELP3 genetic variations may be associated with an increased risk for neurodegeneration and motor neuron diseases. {ECO:0000303|PubMed:18996918}.	central nervous system development [GO:0007417]; histone H3 acetylation [GO:0043966]; histone H4 acetylation [GO:0043967]; neuron migration [GO:0001764]; positive regulation of cell migration [GO:0030335]; regulation of transcription from RNA polymerase II promoter [GO:0006357]; transcription elongation from RNA polymerase II promoter [GO:0006368]

28	P18074	ERCC2_HUMAN	ERCC2 XPD XPDC	TFIIH basal transcription factor complex helicase XPD subunit (EC 3.6.4.12) (Basic transcription factor 2 80 kDa subunit) (BTF2 p80) (CXPD) (DNA excision repair protein ERCC-2) (DNA repair protein complementing XP-D cells) (TFIIH basal transcription factor complex 80 kDa subunit) (TFIIH 80 kDa subunit) (TFIIH p80) (Xeroderma pigmentosum group D-complementing protein)	C116-C134-C155-C190	Fe_4S_4	Structural - Regulatory	3.6.4.12	Cytoplasm, Nucleus	No	DISEASE: Xeroderma pigmentosum complementation group D (XP-D) [MIM:278730]: An autosomal recessive pigmentary skin disorder characterized by solar hypersensitivity of the skin, high predisposition for developing cancers on areas exposed to sunlight and, in some cases, neurological abnormalities. The skin develops marked freckling and other pigmentation abnormalities. Some XP-D patients present features of Cockayne syndrome, including cachectic dwarfism, pigmentary retinopathy, ataxia, decreased nerve conduction velocities. The phenotype combining xeroderma pigmentosum and Cockayne syndrome traits is referred to as XP-CS complex. {ECO:0000269|PubMed:10447254, ECO:0000269|PubMed:11709541, ECO:0000269|PubMed:15494306, ECO:0000269|PubMed:7585650, ECO:0000269|PubMed:7825573, ECO:0000269|PubMed:7849702, ECO:0000269|PubMed:9101292}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Trichothiodystrophy 1, photosensitive (TTD1) [MIM:601675]: A form of trichothiodystrophy, an autosomal recessive disease characterized by sulfur-deficient brittle hair and multisystem variable abnormalities. The spectrum of clinical features varies from mild disease with only hair involvement to severe disease with cutaneous, neurologic and profound developmental defects. Ichthyosis, intellectual and developmental disabilities, decreased fertility, abnormal characteristics at birth, ocular abnormalities, short stature, and infections are common manifestations. There are both photosensitive and non-photosensitive forms of the disorder. TTD1 patients manifest cutaneous photosensitivity. {ECO:0000269|PubMed:11242112, ECO:0000269|PubMed:7920640, ECO:0000269|PubMed:8571952, ECO:0000269|PubMed:9195225, ECO:0000269|PubMed:9238033, ECO:0000269|PubMed:9758621}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Cerebro-oculo-facio-skeletal syndrome 2 (COFS2) [MIM:610756]: A disorder of prenatal onset characterized by microcephaly, congenital cataracts, facial dysmorphism, neurogenic arthrogryposis, growth failure and severe psychomotor retardation. COFS is considered to be part of the nucleotide-excision repair disorders spectrum that include also xeroderma pigmentosum, trichothiodystrophy and Cockayne syndrome. {ECO:0000269|PubMed:11443545}. Note=The disease is caused by mutations affecting the gene represented in this entry.	7-methylguanosine mRNA capping [GO:0006370]; aging [GO:0007568]; apoptotic process [GO:0006915]; bone mineralization [GO:0030282]; cell proliferation [GO:0008283]; central nervous system myelin formation [GO:0032289]; chromosome segregation [GO:0007059]; embryonic cleavage [GO:0040016]; embryonic organ development [GO:0048568]; erythrocyte maturation [GO:0043249]; extracellular matrix organization [GO:0030198]; global genome nucleotide-excision repair [GO:0070911]; hair cell differentiation [GO:0035315]; hair follicle maturation [GO:0048820]; hematopoietic stem cell differentiation [GO:0060218]; in utero embryonic development [GO:0001701]; multicellular organism growth [GO:0035264]; nucleotide-excision repair [GO:0006289]; nucleotide-excision repair, DNA duplex unwinding [GO:0000717]; nucleotide-excision repair, DNA incision [GO:0033683]; nucleotide-excision repair, DNA incision, 3'-to lesion [GO:0006295]; nucleotide-excision repair, DNA incision, 5'-to lesion [GO:0006296]; nucleotide-excision repair, preincision complex assembly [GO:0006294]; nucleotide-excision repair, preincision complex stabilization [GO:0006293]; positive regulation of DNA binding [GO:0043388]; positive regulation of transcription, DNA-templated [GO:0045893]; positive regulation of transcription from RNA polymerase II promoter [GO:0045944]; post-embryonic development [GO:0009791]; protein phosphorylation [GO:0006468]; regulation of mitotic cell cycle phase transition [GO:1901990]; response to hypoxia [GO:0001666]; response to oxidative stress [GO:0006979]; spinal cord development [GO:0021510]; termination of RNA polymerase I transcription [GO:0006363]; transcription-coupled nucleotide-excision repair [GO:0006283]; transcription elongation from RNA polymerase II promoter [GO:0006368]; transcription elongation from RNA polymerase I promoter [GO:0006362]; transcription from RNA polymerase II promoter [GO:0006366]; transcription initiation from RNA polymerase II promoter [GO:0006367]; transcription initiation from RNA polymerase I promoter [GO:0006361]; UV protection [GO:0009650]; viral process [GO:0016032]
29	Q16134	ETFD_HUMAN	ETFDH	Electron transfer flavoprotein-ubiquinone oxidoreductase, mitochondrial (ETF-QO) (ETF-ubiquinone oxidoreductase) (EC 1.5.5.1) (Electron-transferring-flavoprotein dehydrogenase) (ETF dehydrogenase)	C561-C586-C589-C592	Fe_4S_4	Electron transfer	1.5.5.1	Mitochondrion	Yes	DISEASE: Glutaric aciduria 2C (GA2C) [MIM:231680]: An autosomal recessively inherited disorder of fatty acid, amino acid, and choline metabolism. It is characterized by multiple acyl-CoA dehydrogenase deficiencies resulting in large excretion not only of glutaric acid, but also of lactic, ethylmalonic, butyric, isobutyric, 2-methyl-butyric, and isovaleric acids. {ECO:0000269|PubMed:12359134, ECO:0000269|PubMed:12815589, ECO:0000269|PubMed:16527485, ECO:0000269|PubMed:17412732, ECO:0000269|PubMed:19249206, ECO:0000269|PubMed:20370797}. Note=The disease is caused by mutations affecting the gene represented in this entry.	electron transport chain [GO:0022900]; fatty acid beta-oxidation using acyl-CoA dehydrogenase [GO:0033539]; respiratory electron transport chain [GO:0022904]; response to oxidative stress [GO:0006979]

30	Q9BX63	FANCJ_HUMAN	BRIP1 BACH1 FANCJ	Fanconi anemia group J protein (Protein FACJ) (EC 3.6.4.13) (ATP-dependent RNA helicase BRIP1) (BRCA1-associated C-terminal helicase 1) (BRCA1-interacting protein C-terminal helicase 1) (BRCA1-interacting protein 1)	C283-C298-C310-C350	Fe_4S_4	Structural - Regulatory	3.6.4.13	Nucleus	No	DISEASE: Breast cancer (BC) [MIM:114480]: A common malignancy originating from breast epithelial tissue. Breast neoplasms can be distinguished by their histologic pattern. Invasive ductal carcinoma is by far the most common type. Breast cancer is etiologically and genetically heterogeneous. Important genetic factors have been indicated by familial occurrence and bilateral involvement. Mutations at more than one locus can be involved in different families or even in the same case. {ECO:0000269|PubMed:11301010, ECO:0000269|PubMed:14983014}. Note=Disease susceptibility is associated with variations affecting the gene represented in this entry.; DISEASE: Fanconi anemia complementation group J (FANCJ) [MIM:609054]: A disorder affecting all bone marrow elements and resulting in anemia, leukopenia and thrombopenia. It is associated with cardiac, renal and limb malformations, dermal pigmentary changes, and a predisposition to the development of malignancies. At the cellular level it is associated with hypersensitivity to DNA-damaging agents, chromosomal instability (increased chromosome breakage) and defective DNA repair. {ECO:0000269|PubMed:16116423, ECO:0000269|PubMed:16116424, ECO:0000269|PubMed:20639400}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cellular response to angiotensin [GO:1904385]; cellular response to hypoxia [GO:0071456]; cellular response to vitamin [GO:0071295]; chiasma assembly [GO:0051026]; DNA damage checkpoint [GO:0000077]; DNA replication [GO:0006260]; DNA synthesis involved in DNA repair [GO:0000731]; double-strand break repair [GO:0006302]; double-strand break repair involved in meiotic recombination [GO:1990918]; meiotic DNA double-strand break processing involved in reciprocal meiotic recombination [GO:0010705]; negative regulation of cell proliferation [GO:0008285]; negative regulation of gene expression [GO:0010629]; regulation of signal transduction by p53 class mediator [GO:1901796]; regulation of transcription from RNA polymerase II promoter [GO:0006357]; response to toxic substance [GO:0009636]; seminiferous tubule development [GO:0072520]; spermatid development [GO:0007286]; spermatogonial cell division [GO:0007284]; strand displacement [GO:0000732]
31	Q6P4F2	FDX2_HUMAN	FDX2 FDX1L	Ferredoxin-2, mitochondrial (Adrenodoxin-like protein) (Ferredoxin-1-like protein)	C105-C111-C114-C151	Fe_2S_2	Substrate - biogenesis		Mitochondrion	No		C21-steroid hormone biosynthetic process [GO:0006700]; small molecule metabolic process [GO:0044281]; sterol metabolic process [GO:0016125]
32	P80404	GABT_HUMAN	ABAT GABAT	4-aminobutyrate aminotransferase, mitochondrial (EC 2.6.1.19) ((S)-3-amino-2-methylpropionate transaminase) (EC 2.6.1.22) (GABA aminotransferase) (GABA-AT) (Gamma-amino-N-butyrate transaminase) (GABA transaminase) (GABA-T) (L-AIBAT)	C163-C166	Fe_2S_2 per homodimer	Unknown	2.6.1.19; 2.6.1.22	Mitochondrion	No	DISEASE: GABA transaminase deficiency (GABATD) [MIM:613163]: An enzymatic deficiency resulting in psychomotor retardation, hypotonia, hyperreflexia, lethargy, refractory seizures, and EEG abnormalities. {ECO:0000269|PubMed:10407778}. Note=The disease is caused by mutations affecting the gene represented in this entry.	aging [GO:0007568]; behavioral response to cocaine [GO:0048148]; cerebellum development [GO:0021549]; copulation [GO:0007620]; exploration behavior [GO:0035640]; gamma-aminobutyric acid biosynthetic process [GO:0009449]; gamma-aminobutyric acid catabolic process [GO:0009450]; locomotory behavior [GO:0007626]; negative regulation of blood pressure [GO:0045776]; negative regulation of dopamine secretion [GO:0033602]; negative regulation of gamma-aminobutyric acid secretion [GO:0014053]; negative regulation of platelet aggregation [GO:0090331]; neurotransmitter catabolic process [GO:0042135]; positive regulation of aspartate secretion [GO:1904450]; positive regulation of dopamine metabolic process [GO:0045964]; positive regulation of heat generation [GO:0031652]; positive regulation of inhibitory postsynaptic potential [GO:0097151]; positive regulation of insulin secretion [GO:0032024]; positive regulation of prolactin secretion [GO:1902722]; positive regulation of uterine smooth muscle contraction [GO:0070474]; response to drug [GO:0042493]; response to ethanol [GO:0045471]; response to hypoxia [GO:0001666]; response to iron ion [GO:0010039]; response to nicotine [GO:0035094]
33	Q9NS18	GLRX2_HUMAN	GLRX2 GRX2 CGI-133	Glutaredoxin-2, mitochondrial	C77	Fe_2S_2 per homodimer	Substrate - biosinthesis		Mitochondrion	No		aging [GO:0007568]; apoptotic process [GO:0006915]; cell differentiation [GO:0030154]; cell redox homeostasis [GO:0045454]; cellular response to superoxide [GO:0071451]; DNA protection [GO:0042262]; glutathione metabolic process [GO:0006749]; regulation of signal transduction [GO:0009966]; regulation of transcription, DNA-templated [GO:0006355]; response to hydrogen peroxide [GO:0042542]; response to organic substance [GO:0010033]; response to redox state [GO:0051775]; response to temperature stimulus [GO:0009266]
34	O76003	GLRX3_HUMAN	GLRX3 PICOT TXNL2 HUSSY-22	Glutaredoxin-3 (PKC-interacting cousin of thioredoxin) (PICOT) (PKC-theta-interacting protein) (PKCq-interacting protein) (Thioredoxin-like protein 2)	C159-C261	Fe_2S_2 shared with partner	Substrate - biosinthesis		Cytoplasm, Cell membrane	Yes		[2Fe-2S] cluster assembly [GO:0044571]; cell redox homeostasis [GO:0045454]; negative regulation of cardiac muscle hypertrophy [GO:0010614]; protein maturation by iron-sulfur cluster transfer [GO:0097428]; regulation of the force of heart contraction [GO:0002026]

35	Q86SX6	GLRX5_HUMAN	GLRX5 C14orf87	Glutaredoxin-related protein 5, mitochondrial (Monothiol glutaredoxin-5)	C67	Fe_2S_2	Substrate - biogenesis		Mitochondrion	No	DISEASE: Anemia, sideroblastic, 3, pyridoxine-refractory (SIDBA3) [MIM:616860]: A form of sideroblastic anemia, a bone marrow disorder defined by the presence of pathologic iron deposits in erythroblast mitochondria. Sideroblastic anemia is characterized by anemia of varying severity, hypochromic peripheral erythrocytes, systemic iron overload secondary to chronic ineffective erythropoiesis, and the presence of bone marrow ringed sideroblasts. Sideroblasts are characterized by iron-loaded mitochondria clustered around the nucleus. SIDBA3 is refractory to treatment with vitamin B6, while iron chelation therapy may result in clinical improvement. SIDBA3 inheritance is autosomal recessive. {ECO:0000269|PubMed:17485548, ECO:0000269|PubMed:20364084, ECO:0000269|PubMed:25342667, ECO:0000269|PubMed:26100117}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Spasticity, childhood-onset, with hyperglycinemia (SPAHGC) [MIM:616859]: An autosomal recessive disorder characterized by childhood-onset of spasticity, spinal lesions, leukodystrophy, optic atrophy in some patients, non-ketotic hyperglycinemia, and defective enzymatic glycine cleavage. Glycine levels in the cerebrospinal fluid are mildly increased in some but not all patients. The increase is less pronounced than in patients with classic non-ketotic hyperglycinemia. {ECO:0000269|PubMed:24334290}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cell redox homeostasis [GO:0045454]; hemopoiesis [GO:0030097]; protein lipoylation [GO:0009249]; small molecule metabolic process [GO:0044281]
36	A8MXD5	GRCR1_HUMAN	GRXCR1 DFNB25	Glutaredoxin domain-containing cysteine-rich protein 1	C156	Fe_2S_2 (predicted)	Unknown		Unknown	No	DISEASE: Deafness, autosomal recessive, 25 (DFNB25) [MIM:613285]: A form of non-syndromic sensorineural deafness characterized by moderate to severe or profound hearing loss which is progressive in some individuals but not in others. Speech development is impaired in some but not all affected individuals, and vestibular dysfunction is observed in some affected individuals. Sensorineural deafness results from damage to the neural receptors of the inner ear, the nerve pathways to the brain, or the area of the brain that receives sound information. {ECO:0000269|PubMed:20137774, ECO:0000269|PubMed:20137778}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cell redox homeostasis [GO:0045454]; inner ear receptor cell development [GO:0060119]; inner ear receptor stereocilium organization [GO:0060122]; negative regulation of phosphatase activity [GO:0010923]; sensory perception of sound [GO:0007605]; vestibular receptor cell development [GO:0060118]
37	P22830	HEMH_HUMAN	FECH	Ferrochelatase, mitochondrial (EC 4.99.1.1) (Heme synthase) (Protoheme ferro-lyase)	C196-C403-C406-C411	Fe_2S_2	Regulatory	4.99.1.1	Mitochondrion	Yes	DISEASE: Erythropoietic protoporphyria (EPP) [MIM:177000]: A form of porphyria. Porphyrias are inherited defects in the biosynthesis of heme, resulting in the accumulation and increased excretion of porphyrins or porphyrin precursors. They are classified as erythropoietic or hepatic, depending on whether the enzyme deficiency occurs in red blood cells or in the liver. Erythropoietic protoporphyria is marked by excessive protoporphyrin in erythrocytes, plasma, liver and feces, and by widely varying photosensitive skin changes ranging from a burning or pruritic sensation to erythema, edema and wheals. {ECO:0000269|PubMed:10942404, ECO:0000269|PubMed:11375302, ECO:0000269|PubMed:12063482, ECO:0000269|PubMed:12601550, ECO:0000269|PubMed:1376018, ECO:0000269|PubMed:15286165, ECO:0000269|PubMed:17196862, ECO:0000269|PubMed:1755842, ECO:0000269|PubMed:7910885, ECO:0000269|PubMed:8757534, ECO:0000269|PubMed:9211198, ECO:0000269|PubMed:9585598, ECO:0000269|PubMed:9740232}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cellular response to dexamethasone stimulus [GO:0071549]; generation of precursor metabolites and energy [GO:0006091]; heme biosynthetic process [GO:0006783]; protoporphyrinogen IX metabolic process [GO:0046501]; response to arsenic-containing substance [GO:0046685]; response to drug [GO:0042493]; response to ethanol [GO:0045471]; response to insecticide [GO:0017085]; response to lead ion [GO:0010288]; response to light stimulus [GO:0009416]; response to methylmercury [GO:0051597]; response to platinum ion [GO:0070541]
38	P48200	IREB2_HUMAN	IREB2	Iron-responsive element-binding protein 2 (IRE-BP 2) (Iron regulatory protein 2) (IRP2)	C512-C578-C581	Fe_4S_4	Substrate - sensor		Cytoplasm	No		cellular iron ion homeostasis [GO:0006879]; iron ion transport [GO:0006826]; metabolic process [GO:0008152]
39	Q9BUE6	ISCA1_HUMAN	ISCA1 HBLD2 GK004	Iron-sulfur cluster assembly 1 homolog, mitochondrial (HESB-like domain-containing protein 2) (Iron-sulfur assembly protein IscA) (hIscA)	C57-C121-C123	Fe_2S_2/Fe_4S_4	Substrate - biogenesis		Mitochondrion	No	DISEASE: Multiple mitochondrial dysfunctions syndrome 5 (MMDS5) [MIM:617613]: An autosomal recessive, severe disorder characterized by early onset neurological deterioration, seizures, cerebral and cerebellar leukodystrophy, dysmyelination, cortical migrational abnormalities, lactic acidosis and early demise. {ECO:0000269|PubMed:28356563}. Note=The disease is caused by mutations affecting the gene represented in this entry.	iron-sulfur cluster assembly [GO:0016226]; protein maturation by iron-sulfur cluster transfer [GO:0097428]; small molecule metabolic process [GO:0044281]

40	Q86U28	ISCA2_HUMAN	ISCA2 HBLD1	Iron-sulfur cluster assembly 2 homolog, mitochondrial (HESB-like domain-containing protein 1)	C79-C144-C146	Fe_2S_2/Fe_4S_4	Substrate - biogenesis		Mitochondrion	No	DISEASE: Multiple mitochondrial dysfunctions syndrome 4 (MMDS4) [MIM:616370]: A severe disorder of systemic energy metabolism, resulting in weakness, respiratory failure, lack of neurologic development, lactic acidosis, hyperglycinemia and early death. {ECO:0000269|PubMed:25539947}. Note=The disease is caused by mutations affecting the gene represented in this entry.	iron-sulfur cluster assembly [GO:0016226]; protein maturation [GO:0051604]; protein maturation by iron-sulfur cluster transfer [GO:0097428]; small molecule metabolic process [GO:0044281]
41	Q9H1K1	ISCU_HUMAN	ISCU NIFUN	Iron-sulfur cluster assembly enzyme ISCU, mitochondrial (NifU-like N-terminal domain-containing protein) (NifU-like protein)	C69-C95-H137-C138	Fe_2S_2	Substrate - biogenesis		Mitochondrion	No	DISEASE: Myopathy with exercise intolerance Swedish type (MEIS) [MIM:255125]: Autosomal recessive metabolic disease characterized by lifelong severe exercise intolerance, in which minor exertion causes fatigue of active muscles, shortness of breath, and cardiac palpitations in association with lactic acidosis. The biochemical phenotype is characterized by a deficiency in mitochondrial iron-sulfur proteins and impaired muscle oxidative metabolism. {ECO:0000269|PubMed:18304497}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cellular iron ion homeostasis [GO:0006879]; iron-sulfur cluster assembly [GO:0016226]; protein maturation by iron-sulfur cluster transfer [GO:0097428]; small molecule metabolic process [GO:0044281]
42	O43766	LIAS_HUMAN	LIAS LAS HUSSY-01	Lipoyl synthase, mitochondrial (EC 2.8.1.8) (Lipoate synthase) (LS) (Lip-syn) (Lipoic acid synthase)	C106-C111-C117; C137-C141-C144	2 × Fe_4S_4	Electron transfer, Catalytic	2.8.1.8	Mitochondrion	No	DISEASE: Hyperglycinemia, lactic acidosis, and seizures (HGCLAS) [MIM:614462]: An enzymatic defect resulting in an autosomal recessive disorder of mitochondrial metabolism. It is characterized by early-onset lactic acidosis, severe encephalomyopathy, and a pyruvate oxidation defect. Affected individuals have neonatal-onset epilepsy, poor growth, psychomotor retardation, muscular hypotonia, lactic acidosis, and elevated glycine concentration in plasma and urine. {ECO:0000269|PubMed:22152680}. Note=The disease is caused by mutations affecting the gene represented in this entry.	cellular nitrogen compound metabolic process [GO:0034641]; inflammatory response [GO:0006954]; lipoate biosynthetic process [GO:0009107]; neural tube closure [GO:0001843]; protein lipoylation [GO:0009249]; response to lipopolysaccharide [GO:0032496]; response to oxidative stress [GO:0006979]
43	Q9NZB8	MOCS1_HUMAN	MOCS1 MIG11	Molybdenum cofactor biosynthesis protein 1 (Cell migration-inducing gene 11 protein) (Molybdenum cofactor synthesis-step 1 protein A-B) [Includes: GTP 3',8-cyclase (EC 4.1.99.22) (Molybdenum cofactor biosynthesis protein A); Cyclic pyranopterin monophosphate synthase (EC 4.6.1.17) (Molybdenum cofactor biosynthesis protein C)]	C80-C84-C87; C312-C315-C329	2 × Fe_4S_4	Catalytic, Structural	4.1.99.22; 4.6.1.17	Unknown	No	DISEASE: Molybdenum cofactor deficiency, complementation group A (MOCODA) [MIM:252150]: An autosomal recessive metabolic disorder leading to the pleiotropic loss of molybdoenzyme activities. It is clinically characterized by onset in infancy of poor feeding, intractable seizures, severe psychomotor retardation, and death in early childhood in most patients. {ECO:0000269|PubMed:12754701, ECO:0000269|PubMed:16021469, ECO:0000269|PubMed:9731530, ECO:0000269|PubMed:9921896}. Note=The disease is caused by mutations affecting the gene represented in this entry.	molybdopterin cofactor biosynthetic process [GO:0032324]; Mo-molybdopterin cofactor biosynthetic process [GO:0006777]
44	Q9UIF7	MUTYH_HUMAN	MUTYH MYH	Adenine DNA glycosylase (EC 3.2.2.-) (MutY homolog) (hMYH)	C287-C294-C297-C303	Fe_4S_4	Structural - Regulatory	3.2.2.-	Mitochondrion, Nucleus	No	DISEASE: Familial adenomatous polyposis 2 (FAP2) [MIM:608456]: A condition characterized by the development of multiple colorectal adenomatous polyps, benign neoplasms derived from glandular epithelium. Some affected individuals may develop colorectal carcinoma. {ECO:0000269|PubMed:11818965, ECO:0000269|PubMed:12606733, ECO:0000269|PubMed:12853198, ECO:0000269|PubMed:15366000, ECO:0000269|PubMed:16134147, ECO:0000269|PubMed:16287072, ECO:0000269|PubMed:16557584, ECO:0000269|PubMed:16941501, ECO:0000269|PubMed:18091433, ECO:0000269|PubMed:18515411, ECO:0000269|PubMed:19953527, ECO:0000269|PubMed:20418187, ECO:0000269|PubMed:20848659, ECO:0000269|PubMed:25820570, ECO:0000269|PubMed:26694661}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Gastric cancer (GASC) [MIM:613659]: A malignant disease which starts in the stomach, can spread to the esophagus or the small intestine, and can extend through the stomach wall to nearby lymph nodes and organs. It also can metastasize to other parts of the body. The term gastric cancer or gastric carcinoma refers to adenocarcinoma of the stomach that accounts for most of all gastric malignant tumors. Two main histologic types are recognized, diffuse type and intestinal type carcinomas. Diffuse tumors are poorly differentiated infiltrating lesions, resulting in thickening of the stomach. In contrast, intestinal tumors are usually exophytic, often ulcerating, and associated with intestinal metaplasia of the stomach, most often observed in sporadic disease. {ECO:0000269|PubMed:15273732, ECO:0000269|PubMed:25820570}. Note=The gene represented in this entry may be involved in disease pathogenesis. Somatic mutations contribute to the development of a sub-set of sporadic gastric cancers in carriers of Helicobacter pylori (PubMed:15273732). {ECO:0000269|PubMed:15273732}.	depurination [GO:0045007]; DNA repair [GO:0006281]; mismatch repair [GO:0006298]

45	Q9UHQ1	NARF_HUMAN	NARF	Nuclear prelamin A recognition factor (Iron-only hydrogenase-like protein 2) (IOP2)	C172-C228-C374-C378	2 × Fe_4S_4	Unknown		Nucleus	No		
46	Q9H6Q4	NARFL_HUMAN	NARFL PRN	Cytosolic Fe-S cluster assembly factor NARFL (Iron-only hydrogenase-like protein 1) (IOP1) (Nuclear prelamin A recognition factor-like protein) (Protein related to Narf)	C24-C71-C74-C77; C190-C246-C395-C399	2 × Fe_4S_4	Substrate - biogenesis		Unknown	No		hematopoietic progenitor cell differentiation [GO:0002244]; iron-sulfur cluster assembly [GO:0016226]; oxygen homeostasis [GO:0032364]; regulation of gene expression [GO:0010468]; response to hypoxia [GO:0001666]
47	P28331	NDUS1_HUMAN	NDUFS1	NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial (EC 1.6.5.3) (EC 1.6.99.3) (Complex I-75kD) (CI-75kD)	C64-C75-C78-C92; H124-C128-C131-C137; C176-C179-C182-C226	2 × Fe_4S_4, Fe_2S_2	Electron transfer	1.6.5.3; 1.6.99.3	Mitochondrion	Yes	DISEASE: Mitochondrial complex I deficiency (MT-C1D) [MIM:252010]: A disorder of the mitochondrial respiratory chain that causes a wide range of clinical manifestations from lethal neonatal disease to adult-onset neurodegenerative disorders. Phenotypes include macrocephaly with progressive leukodystrophy, non-specific encephalopathy, cardiomyopathy, myopathy, liver disease, Leigh syndrome, Leber hereditary optic neuropathy, and some forms of Parkinson disease. {ECO:0000269|PubMed:11349233}. Note=The disease is caused by mutations affecting the gene represented in this entry.	apoptotic mitochondrial changes [GO:0008637]; ATP metabolic process [GO:0046034]; cellular respiration [GO:0045333]; mitochondrial electron transport, NADH to ubiquinone [GO:0006120]; mitochondrial respiratory chain complex I assembly [GO:0032981]; reactive oxygen species metabolic process [GO:0072593]; regulation of mitochondrial membrane potential [GO:0051881]
48	O75306	NDUS2_HUMAN	NDUFS2	NADH dehydrogenase [ubiquinone] iron-sulfur protein 2, mitochondrial (EC 1.6.5.3) (EC 1.6.99.3) (Complex I-49kD) (CI-49kD) (NADH-ubiquinone oxidoreductase 49 kDa subunit)	C326-C332-C347	Fe_4S_4	Electron transfer	1.6.5.3; 1.6.99.3	Mitochondrion	Yes	DISEASE: Mitochondrial complex I deficiency (MT-C1D) [MIM:252010]: A disorder of the mitochondrial respiratory chain that causes a wide range of clinical manifestations from lethal neonatal disease to adult-onset neurodegenerative disorders. Phenotypes include macrocephaly with progressive leukodystrophy, non-specific encephalopathy, cardiomyopathy, myopathy, liver disease, Leigh syndrome, Leber hereditary optic neuropathy, and some forms of Parkinson disease. {ECO:0000269|PubMed:11220739}. Note=The disease is caused by mutations affecting the gene represented in this entry.	mitochondrial ATP synthesis coupled electron transport [GO:0042775]; mitochondrial electron transport, NADH to ubiquinone [GO:0006120]; mitochondrial respiratory chain complex I assembly [GO:0032981]; response to oxidative stress [GO:0006979]
49	O75251	NDUS7_HUMAN	NDUFS7	NADH dehydrogenase [ubiquinone] iron-sulfur protein 7, mitochondrial (EC 1.6.5.3) (EC 1.6.99.3) (Complex I-20kD) (CI-20kD) (NADH-ubiquinone oxidoreductase 20 kDa subunit) (PSST subunit)	C88-C89-C153-C183	Fe_4S_4	Electron transfer	1.6.5.3; 1.6.99.3	Mitochondrion	No	DISEASE: Leigh syndrome (LS) [MIM:256000]: An early-onset progressive neurodegenerative disorder characterized by the presence of focal, bilateral lesions in one or more areas of the central nervous system including the brainstem, thalamus, basal ganglia, cerebellum and spinal cord. Clinical features depend on which areas of the central nervous system are involved and include subacute onset of psychomotor retardation, hypotonia, ataxia, weakness, vision loss, eye movement abnormalities, seizures, and dysphagia. {ECO:0000269|PubMed:10360771}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Mitochondrial complex I deficiency (MT-C1D) [MIM:252010]: A disorder of the mitochondrial respiratory chain that causes a wide range of clinical manifestations from lethal neonatal disease to adult-onset neurodegenerative disorders. Phenotypes include macrocephaly with progressive leukodystrophy, non-specific encephalopathy, cardiomyopathy, myopathy, liver disease, Leigh syndrome, Leber hereditary optic neuropathy, and some forms of Parkinson disease. {ECO:0000269|PubMed:10330338}. Note=The disease is caused by mutations affecting the gene represented in this entry.	mitochondrial electron transport, NADH to ubiquinone [GO:0006120]; mitochondrial respiratory chain complex I assembly [GO:0032981]
50	O00217	NDUS8_HUMAN	NDUFS8	NADH dehydrogenase [ubiquinone] iron-sulfur protein 8, mitochondrial (EC 1.6.5.3) (EC 1.6.99.3) (Complex I-23kD) (CI-23kD) (NADH-ubiquinone oxidoreductase 23 kDa subunit) (TYKY subunit)	C111-C114-C117-C160; C121-C150-C153-C156	2 × Fe_4S_4	Electron transfer	1.6.5.3; 1.6.99.3	Mitochondrion	No	DISEASE: Leigh syndrome (LS) [MIM:256000]: An early-onset progressive neurodegenerative disorder characterized by the presence of focal, bilateral lesions in one or more areas of the central nervous system including the brainstem, thalamus, basal ganglia, cerebellum and spinal cord. Clinical features depend on which areas of the central nervous system are involved and include subacute onset of psychomotor retardation, hypotonia, ataxia, weakness, vision loss, eye movement abnormalities, seizures, and dysphagia. {ECO:0000269|PubMed:9837812}. Note=The disease is caused by mutations affecting the gene represented in this entry.	mitochondrial electron transport, NADH to ubiquinone [GO:0006120]; mitochondrial respiratory chain complex I assembly [GO:0032981]; response to oxidative stress [GO:0006979]

51	P49821	NDUV1_HUMAN	NDUFV1 UQOR1	NADH dehydrogenase [ubiquinone] flavoprotein 1, mitochondrial (EC 1.6.5.3) (EC 1.6.99.3) (Complex I-51kD) (CI-51kD) (NADH dehydrogenase flavoprotein 1) (NADH-ubiquinone oxidoreductase 51 kDa subunit)	C379-C382-C385-C425	Fe_4S_4	Electron transfer	1.6.5.3; 1.6.99.3	Mitochondrion	Yes	DISEASE: Leigh syndrome (LS) [MIM:256000]: An early-onset progressive neurodegenerative disorder characterized by the presence of focal, bilateral lesions in one or more areas of the central nervous system including the brainstem, thalamus, basal ganglia, cerebellum and spinal cord. Clinical features depend on which areas of the central nervous system are involved and include subacute onset of psychomotor retardation, hypotonia, ataxia, weakness, vision loss, eye movement abnormalities, seizures, and dysphagia. {ECO:0000269|PubMed:10080174}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Mitochondrial complex I deficiency (MT-C1D) [MIM:252010]: A disorder of the mitochondrial respiratory chain that causes a wide range of clinical manifestations from lethal neonatal disease to adult-onset neurodegenerative disorders. Phenotypes include macrocephaly with progressive leukodystrophy, non-specific encephalopathy, cardiomyopathy, myopathy, liver disease, Leigh syndrome, Leber hereditary optic neuropathy, and some forms of Parkinson disease. {ECO:0000269|PubMed:10080174, ECO:0000269|PubMed:11349233}. Note=The disease is caused by mutations affecting the gene represented in this entry.	mitochondrial ATP synthesis coupled electron transport [GO:0042775]; mitochondrial electron transport, NADH to ubiquinone [GO:0006120]; mitochondrial respiratory chain complex I assembly [GO:0032981]
52	P19404	NDUV2_HUMAN	NDUFV2	NADH dehydrogenase [ubiquinone] flavoprotein 2, mitochondrial (EC 1.6.5.3) (EC 1.6.99.3) (NADH-ubiquinone oxidoreductase 24 kDa subunit)	C135-C140-C176-C180	Fe_2S_2	Electron transfer	1.6.5.3; 1.6.99.3	Mitochondrion	Yes		cardiac muscle tissue development [GO:0048738]; mitochondrial electron transport, NADH to ubiquinone [GO:0006120]; mitochondrial respiratory chain complex I assembly [GO:0032981]; nervous system development [GO:0007399]
53	Q9Y697	NFS1_HUMAN	NFS1 NIFS HUSSY-08	Cysteine desulfurase, mitochondrial (EC 2.8.1.7)	C381	Fe_2S_2	Substrate - biogenesis	2.8.1.7	Mitochondrion	No		[2Fe-2S] cluster assembly [GO:0044571]; iron incorporation into metallo-sulfur cluster [GO:0018283]; molybdopterin cofactor biosynthetic process [GO:0032324]; Mo-molybdopterin cofactor biosynthetic process [GO:0006777]; protein complex assembly [GO:0006461]; small molecule metabolic process [GO:0044281]; sulfur amino acid metabolic process [GO:0000096]
54	Q9UMS0	NFU1_HUMAN	NFU1 HIRIP5 CGI-33	NFU1 iron-sulfur cluster scaffold homolog, mitochondrial (HIRA-interacting protein 5)	C210-C213	Fe_4S_4	Substrate - biogenesis		Cytoplasm, Mitochondrion	No	DISEASE: Multiple mitochondrial dysfunctions syndrome 1 (MMDS1) [MIM:605711]: A severe disorder of systemic energy metabolism, resulting in weakness, respiratory failure, lack of neurologic development, lactic acidosis, hyperglycinemia and early death. Some patients show failure to thrive, pulmonary hypertension, hypotonia and irritability. Biochemical features include severe combined deficiency of the 2-oxoacid dehydrogenases, defective lipoic acid synthesis and reduction in activity of mitochondrial respiratory chain complexes. {ECO:0000269|PubMed:21944046, ECO:0000269|PubMed:22077971, ECO:0000269|PubMed:25918518, ECO:0000269|PubMed:28161430, ECO:0000269|PubMed:28906594}. Note=The disease is caused by mutations affecting the gene represented in this entry.	iron-sulfur cluster assembly [GO:0016226]
55	P78549	NTH_HUMAN	NTHL1 NTH1 OCTS3	Endonuclease III-like protein 1 (hNTH1) (EC 3.2.2.-) (EC 4.2.99.18) (Bifunctional DNA N-glycosylase/DNA-(apurinic or apyrimidinic site) lyase) (DNA glycosylase/AP lyase)	C290-C297-C300-C306	Fe_4S_4	Structural - Regulatory	3.2.2.-; 4.2.99.18	Mitochondrion, Nucleus	No	DISEASE: Familial adenomatous polyposis 3 (FAP3) [MIM:616415]: A form of familial adenomatous polyposis, a condition characterized by the development of multiple colorectal adenomatous polyps, benign neoplasms derived from glandular epithelium. Some affected individuals may develop colorectal carcinoma. {ECO:0000269|PubMed:25938944}. Note=The disease is caused by mutations affecting the gene represented in this entry.	base-excision repair, AP site formation [GO:0006285]; depyrimidination [GO:0045008]; nucleotide-excision repair, DNA incision, 5'-to lesion [GO:0006296]
56	P53384	NUBP1_HUMAN	NUBP1 NBP NBP1	Cytosolic Fe-S cluster assembly factor NUBP1 (Nucleotide-binding protein 1) (NBP 1)	C8-C22-C25-C31; C235-C238	Fe_4S_4, Fe_4S_4 shared with NUBP2	Substrate - biogenesis		Cytoplasm, Nucleus	No		cell growth [GO:0016049]; cell projection organization [GO:0030030]; cellular iron ion homeostasis [GO:0006879]; centrosome localization [GO:0051642]; iron-sulfur cluster assembly [GO:0016226]; negative regulation of centrosome duplication [GO:0010826]; protein localization to cell cortex [GO:0072697]
57	Q9Y5Y2	NUBP2_HUMAN	NUBP2	Cytosolic Fe-S cluster assembly factor NUBP2 (Nucleotide-binding protein 2) (NBP 2)	C196-C199	Fe_4S_4 shared with NUBP1	Substrate - biogenesis		Cytoplasm, Nucleus	No		cell projection organization [GO:0030030]; iron-sulfur cluster assembly [GO:0016226]
58	Q8TB37	NUBPL_HUMAN	NUBPL C14orf127	Iron-sulfur protein NUBPL (IND1 homolog) (Nucleotide-binding protein-like) (huInd1)	C244-C247	Fe_2S_2/Fe_4S_4	Substrate - biogenesis		Mitochondrion	No	DISEASE: Mitochondrial complex I deficiency (MT-C1D) [MIM:252010]: A disorder of the mitochondrial respiratory chain that causes a wide range of clinical manifestations from lethal neonatal disease to adult-onset neurodegenerative disorders. Phenotypes include macrocephaly with progressive leukodystrophy, non-specific encephalopathy, cardiomyopathy, myopathy, liver disease, Leigh syndrome, Leber hereditary optic neuropathy, and some forms of Parkinson disease. {ECO:0000269|PubMed:20818383, ECO:0000269|PubMed:23553477}. Note=The disease is caused by mutations affecting the gene represented in this entry.	mitochondrial respiratory chain complex I assembly [GO:0032981]; mitochondrion morphogenesis [GO:0070584]

59	P49643	PRI2_HUMAN	PRIM2 PRIM2A	DNA primase large subunit (EC 2.7.7.-) (DNA primase 58 kDa subunit) (p58)	C287-C367-C384-C424	Fe_4S_4	Structural - Regulatory	2.7.7.-	Unknown	No		DNA replication, synthesis of RNA primer [GO:0006269]; DNA replication initiation [GO:0006270]; G1/S transition of mitotic cell cycle [GO:0000082]; telomere maintenance via semi-conservative replication [GO:0032201]
60	Q06203	PUR1_HUMAN	PPAT GPAT	Amidophosphoribosyltransferase (ATase) (EC 2.4.2.14) (Glutamine phosphoribosylpyrophosphate amidotransferase) (GPAT)	C280-C426-C503-C506	Fe_4S_4	Unknown	2.4.2.14	Unknown	No		'de novo' IMP biosynthetic process [GO:0006189]; animal organ regeneration [GO:0031100]; cellular response to drug [GO:0035690]; cellular response to insulin stimulus [GO:0032869]; G1/S transition of mitotic cell cycle [GO:0000082]; glutamine catabolic process [GO:0006543]; kidney development [GO:0001822]; lactation [GO:0007595]; maternal process involved in female pregnancy [GO:0060135]; nucleoside metabolic process [GO:0009116]; protein homotetramerization [GO:0051289]; purine nucleobase biosynthetic process [GO:0009113]; purine nucleotide biosynthetic process [GO:0006164]; purine ribonucleoside monophosphate biosynthetic process [GO:0009168]
61	O60673	REV3L_HUMAN	REV3L POLZ REV3	DNA polymerase zeta catalytic subunit (EC 2.7.7.7) (Protein reversionless 3-like) (REV3-like) (hREV3)	C3086-C3089-C3099-C3104	Fe_4S_4	Structural - Regulatory	2.7.7.7	Nucleus	No		DNA-dependent DNA replication [GO:0006261]; error-prone translesion synthesis [GO:0042276]
62	Q8TAC1	RFESD_HUMAN	RFESD	Rieske domain-containing protein	C57-H59-C80-H83	Fe_2S_2 (predicted)	Unknown		Unknown	No		
63	Q9HA92	RSAD1_HUMAN	RSAD1	Radical S-adenosyl methionine domain-containing protein 1, mitochondrial (EC 1.3.99.-) (Oxygen-independent coproporphyrinogen-III oxidase-like protein RSAD1)	C49-C53-C56	Fe_4S_4	Catalytic	1.3.99.-	Mitochondrion	No		porphyrin-containing compound biosynthetic process [GO:0006779]
64	Q8WXG1	RSAD2_HUMAN	RSAD2 CIG5	Radical S-adenosyl methionine domain-containing protein 2 (Cytomegalovirus-induced gene 5 protein) (Viperin) (Virus inhibitory protein, endoplasmic reticulum-associated, interferon-inducible)	C83-C87-C90	Fe_4S_4	Unknown		Cytoplasm, Endoplasmic reticulum, Golgi apparatus, Mitochondrion	Yes		CD4-positive, alpha-beta T cell activation [GO:0035710]; CD4-positive, alpha-beta T cell differentiation [GO:0043367]; defense response to virus [GO:0051607]; negative regulation of protein secretion [GO:0050709]; negative regulation of viral genome replication [GO:0045071]; positive regulation of T-helper 2 cell cytokine production [GO:2000553]; positive regulation of toll-like receptor 7 signaling pathway [GO:0034157]; positive regulation of toll-like receptor 9 signaling pathway [GO:0034165]; response to virus [GO:0009615]; type I interferon signaling pathway [GO:0060337]; viral process [GO:0016032]

65	Q9NZ71	RTEL1_HUMAN	RTEL1 C20orf41 KIAA1088 NHL	Regulator of telomere elongation helicase 1 (EC 3.6.4.12) (Novel helicase-like)	C145-C163-C172-C207	Fe_4S_4	Structural - Regulatory	3.6.4.12	Nucleus	No	DISEASE: Dyskeratosis congenita, autosomal recessive, 5 (DKCB5) [MIM:615190]: A form of dyskeratosis congenita, a rare multisystem disorder caused by defective telomere maintenance. It is characterized by progressive bone marrow failure, and the clinical triad of reticulated skin hyperpigmentation, nail dystrophy, and mucosal leukoplakia. Common but variable features include premature graying, aplastic anemia, low platelets, osteoporosis, pulmonary fibrosis, and liver fibrosis among others. Early mortality is often associated with bone marrow failure, infections, fatal pulmonary complications, or malignancy. DKCB5 is characterized by onset of bone marrow failure and immunodeficiency in early childhood. Most patients also have growth and developmental delay and cerebellar hypoplasia, consistent with a clinical diagnosis of Hoyeraal-Hreidarsson syndrome. {ECO:0000269|PubMed:23329068, ECO:0000269|PubMed:23453664, ECO:0000269|PubMed:23591994, ECO:0000269|PubMed:23959892, ECO:0000269|PubMed:24009516}. Note=The disease is caused by mutations affecting the gene represented in this entry. RTEL1 mutations have also been found in patients with a dyskeratosis congenita-like phenotype consisting of one feature of dyskeratosis congenita and short telomeres, in the absence of the typical DKC diagnostic triad (PubMed:23329068). {ECO:0000269|PubMed:23329068}.; DISEASE: Dyskeratosis congenita, autosomal dominant, 4 (DKCA4) [MIM:615190]: A rare multisystem disorder caused by defective telomere maintenance. It is characterized by progressive bone marrow failure, and the clinical triad of reticulated skin hyperpigmentation, nail dystrophy, and mucosal leukoplakia. Common but variable features include premature graying, aplastic anemia, low platelets, osteoporosis, pulmonary fibrosis, and liver fibrosis among others. Early mortality is often associated with bone marrow failure, infections, fatal pulmonary complications, or malignancy. {ECO:0000269|PubMed:23329068}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Pulmonary fibrosis, and/or bone marrow failure, telomere-related, 3 (PFBMFT3) [MIM:616373]: A disease associated with shortened telomeres. Pulmonary fibrosis is the most common manifestation. Other manifestations include aplastic anemia due to bone marrow failure, hepatic fibrosis, and increased cancer risk, particularly myelodysplastic syndrome and acute myeloid leukemia. Phenotype, age at onset, and severity are determined by telomere length. {ECO:0000269|PubMed:25848748}. Note=The disease is caused by mutations affecting the gene represented in this entry.	DNA duplex unwinding [GO:0032508]; DNA repair [GO:0006281]; mitotic telomere maintenance via semi-conservative replication [GO:1902990]; negative regulation of DNA recombination [GO:0045910]; negative regulation of t-circle formation [GO:1904430]; negative regulation of telomere maintenance in response to DNA damage [GO:1904506]; positive regulation of telomere capping [GO:1904355]; positive regulation of telomere maintenance [GO:0032206]; positive regulation of telomere maintenance via telomere lengthening [GO:1904358]; positive regulation of telomeric loop disassembly [GO:1904535]; regulation of double-strand break repair via homologous recombination [GO:0010569]; replication fork processing [GO:0031297]; strand displacement [GO:0000732]; telomere maintenance [GO:0000723]; telomere maintenance in response to DNA damage [GO:0043247]; telomeric loop disassembly [GO:0090657]

66	P21912	SDHB_HUMAN	SDHB SDH SDH1	Succinate dehydrogenase [ubiquinone] iron-sulfur subunit, mitochondrial (EC 1.3.5.1) (Iron-sulfur subunit of complex II) (Ip)	C93-C98-C101-C113; C186-C189-C192-C253; C196-C243-C249	Fe_2S_2, Fe_3S_4, Fe_4S_4	Electron transfer	1.3.5.1	Mitochondrion	Yes	DISEASE: Pheochromocytoma (PCC) [MIM:171300]: A catecholamine-producing tumor of chromaffin tissue of the adrenal medulla or sympathetic paraganglia. The cardinal symptom, reflecting the increased secretion of epinephrine and norepinephrine, is hypertension, which may be persistent or intermittent. {ECO:0000269|PubMed:11404820, ECO:0000269|PubMed:12000816, ECO:0000269|PubMed:12618761, ECO:0000269|PubMed:14500403, ECO:0000269|PubMed:14974914, ECO:0000269|PubMed:15328326, ECO:0000269|PubMed:17634472}. Note=Disease susceptibility is associated with variations affecting the gene represented in this entry.; DISEASE: Paragangliomas 4 (PGL4) [MIM:115310]: A neural crest tumor usually derived from the chromoreceptor tissue of a paraganglion. Paragangliomas can develop at various body sites, including the head, neck, thorax and abdomen. Most commonly, they are located in the head and neck region, specifically at the carotid bifurcation, the jugular foramen, the vagal nerve, and in the middle ear. {ECO:0000269|PubMed:11404820, ECO:0000269|PubMed:11897817, ECO:0000269|PubMed:14715873, ECO:0000269|PubMed:14974914, ECO:0000269|PubMed:15328326}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Paraganglioma and gastric stromal sarcoma (PGGSS) [MIM:606864]: Gastrointestinal stromal tumors may be sporadic or inherited in an autosomal dominant manner, alone or as a component of a syndrome associated with other tumors, such as in the context of neurofibromatosis type 1 (NF1). Patients have both gastrointestinal stromal tumors and paragangliomas. Susceptibility to the tumors was inherited in an apparently autosomal dominant manner, with incomplete penetrance. {ECO:0000269|PubMed:17804857}. Note=The disease is caused by mutations affecting the gene represented in this entry.; DISEASE: Cowden syndrome 2 (CWS2) [MIM:612359]: A form of Cowden syndrome, a hamartomatous polyposis syndrome with age-related penetrance. Cowden syndrome is characterized by hamartomatous lesions affecting derivatives of ectodermal, mesodermal and endodermal layers, macrocephaly, facial trichilemmomas (benign tumors of the hair follicle infundibulum), acral keratoses, papillomatous papules, and elevated risk for development of several types of malignancy, particularly breast carcinoma in women and thyroid carcinoma in both men and women. Colon cancer and renal cell carcinoma have also been reported. Hamartomas can be found in virtually every organ, but most commonly in the skin, gastrointestinal tract, breast and thyroid. CWS2 inheritance is autosomal dominant. {ECO:0000269|PubMed:18678321}. Note=The disease may be caused by mutations affecting the gene represented in this entry.	aerobic respiration [GO:0009060]; respiratory electron transport chain [GO:0022904]; succinate metabolic process [GO:0006105]; tricarboxylic acid cycle [GO:0006099]
67	Q6NUM6	TYW1B_HUMAN	TYW1B RSAFD2	S-adenosyl-L-methionine-dependent tRNA 4-demethylwyosine synthase (EC 4.1.3.44) (Radical S-adenosyl methionine and flavodoxin domain-containing protein 2) (tRNA wybutosine-synthesizing protein 1 homolog B)	C352-C356-C359	Fe_4S_4	Catalytic	4.1.3.44	Unknown	No		oxidation-reduction process [GO:0055114]; tRNA processing [GO:0008033]
68	P47985	UCRI_HUMAN	UQCRFS1	Cytochrome b-c1 complex subunit Rieske, mitochondrial (EC 1.10.2.2) (Complex III subunit 5) (Cytochrome b-c1 complex subunit 5) (Rieske iron-sulfur protein) (RISP) (Rieske protein UQCRFS1) (Ubiquinol-cytochrome c reductase iron-sulfur subunit) [Cleaved into: Cytochrome b-c1 complex subunit 9 (Su9) (Subunit 9) (8 kDa subunit 9) (Complex III subunit IX) (Cytochrome b-c1 complex subunit 11) (Ubiquinol-cytochrome c reductase 8 kDa protein)]	C217-H219-C236-H239	Fe_2S_2	Electron transfer	1.10.2.2	Mitochondrion	Yes		mitochondrial electron transport, ubiquinol to cytochrome c [GO:0006122]; response to antibiotic [GO:0046677]; response to drug [GO:0042493]; response to hormone [GO:0009725]
69	P0C7P4	UCRIL_HUMAN	UQCRFS1P1 UQCRFSL1	Putative cytochrome b-c1 complex subunit Rieske-like protein 1 (Ubiquinol-cytochrome c reductase Rieske iron-sulfur subunit pseudogene 1)	C226-H228-C231-C245-H248	Fe_2S_2 (predicted)	Unknown		Unknown	No		

70	P47989	XDH_HUMAN	XDH XDHA	Xanthine dehydrogenase/oxidase [Includes: Xanthine dehydrogenase (XD) (EC 1.17.1.4); Xanthine oxidase (XO) (EC 1.17.3.2) (Xanthine oxidoreductase) (XOR)]	C43-C48-C51-C73; C113-C116-C148-C150	2 × Fe_2S_2	Electron transfer	1.17.1.4; 1.17.3.2	Cytoplasm, Extracellular space, Peroxisome	No	DISEASE: Xanthinuria 1 (XAN1) [MIM:278300]: A disorder characterized by excretion of very large amounts of xanthine in the urine and a tendency to form xanthine stones. Uric acid is strikingly diminished in serum and urine. XAN1 is due to isolated xanthine dehydrogenase deficiency. Patients can metabolize allopurinol. {ECO:0000269|PubMed:10844591, ECO:0000269|PubMed:11379872, ECO:0000269|PubMed:14551354, ECO:0000269|PubMed:9153281}. Note=The disease is caused by mutations affecting the gene represented in this entry.	activation of cysteine-type endopeptidase activity involved in apoptotic process [GO:0006919]; lactation [GO:0007595]; negative regulation of endothelial cell differentiation [GO:0045602]; negative regulation of endothelial cell proliferation [GO:0001937]; negative regulation of gene expression [GO:0010629]; negative regulation of protein kinase B signaling [GO:0051898]; negative regulation of protein phosphorylation [GO:0001933]; negative regulation of vascular endothelial growth factor signaling pathway [GO:1900747]; negative regulation of vasculogenesis [GO:2001213]; positive regulation of p38MAPK cascade [GO:1900745]; positive regulation of reactive oxygen species metabolic process [GO:2000379]; purine nucleotide catabolic process [GO:0006195]; xanthine catabolic process [GO:0009115]

5.3 The hMeProt database of human metal-binding proteins

The hMeProt database is a new resource I developed to provide the scientific community with a comprehensive view of the human metalloproteome, i.e., the entire set of metal-binding proteins encoded in the human genome. The hMeProt database integrates data from various biological resources, so as to associate each human metal-binding protein with the largest possible amount of information, and thus facilitate the process of knowledge discovery by the users. Proteins in hMeProt are referenced by the most common identifiers such as UniProt[44] and NCBI accession codes, and are associated with available data on, e.g., cellular localization, GO[74] function, known interactions, metabolic pathways, mutations and associated pathologies, SNPs, tissue expression, etc. To optimize the usage and the effectiveness of hMeProt, I designed and set up a user-friendly web interface which allows one to formulate a wide range of queries, from pre-compiled standard queries to complex, purpose-constructed queries based on Boolean logic. The display of data contained in hMeProt occurs by means of different page templates, each corresponding to, and optimized to convey, a certain type of information. hMeProt can be used to study how variation of human genes encoding metal-binding proteins affects the metal site(s) of those proteins, and how, in turn, this can have a role on cellular function and overall phenotypes. This use case represents an example of how hMeProt can provide insights into the role of metal ions in healthy metabolism, and their impact in the onset and development of human diseases.

5.3.1 Content of the hMeProt database

As described above, the core data stored in the hMeProt database are human metal-binding proteins, which were identified starting from the amino acid sequences in the complete translated human genome. For the identification of such proteins I developed a protocol that combines five different methods, which are presented in detail in the "Methods" section of this thesis. The primary sources of information underlying these methods are (i) the MetalPDB[31] database; (ii) the UniProt[44] database of protein sequences; (iii) the Prosite[73] database of protein domains, families and functional sites [7]; and (iv) the Gene Ontology[74] (GO) framework defining classes used to describe gene function [8]. Currently, the hMeProt database contains 3969 metal-binding proteins, which approximately represent 20% of the whole human proteome, and collectively encompass 11145 metal sites.

When available, hMeProt collects, for each protein entry, a range of additional data retrieved from various resources, including (i) tissue expression, (ii) subcellular location, (iii) function,

(iv) sequence variants (e.g., SNPs, RNA editing events), (v) pathways (molecular interactions, reaction and relation networks), and (vi) pathologies involving the protein. Some of these information could not be obtained for all proteins, either because they are not available, or because they do not apply to all cases (for example, not all proteins are associated with pathologies). On the other hand, hMeProt provides these additional data also for human proteins that are not metal-binding, thereby facilitating comparative analyses between metal-binding and non-metal-binding proteins with similar features. The number of proteins in hMeProt for which the various types of additional information are available is reported in Table 9.

Table 9: *Number of proteins with additional information in the hMeProt database.*

Feature	**Metal-binding proteins**	**Metal-free proteins**	**Total**
Tissue	2.518	9.384	11.902
Subcellular location	3.513	12.854	16.367
Function	3.646	12.526	16.172
Variants	3.549	11.743	15.292
Pathways	980	869	1.849
Pathologies	1.529	1.130	2.659
Drugs	980	869	1.849

5.3.2 hMeProt protein pages

Protein pages are the primary way to present the information contained in the hMeProt database. Each protein page contains the data associated with a human protein (see Figure 18 for example), and is divided into three sections: (i) general information, (ii) metal sites, and (iii) sequence variants outside the metal site.

The first section (general information) is found at the top of the page, where are shown: (i) the UniProt[44] accession code; (ii) the name of the gene encoding the protein; (iii) the sequence length; (iv) the EC number (for enzymes); (v) the subcellular location(s); (vi) the function(s); (vii) the tissue(s) in which the protein is expressed (also indicating the cellular types in which the protein is expressed and the expression level for each cellular type).; (viii) the diseases and (ix) the pathways involving the protein; and (x) the drugs known to interact with the protein.

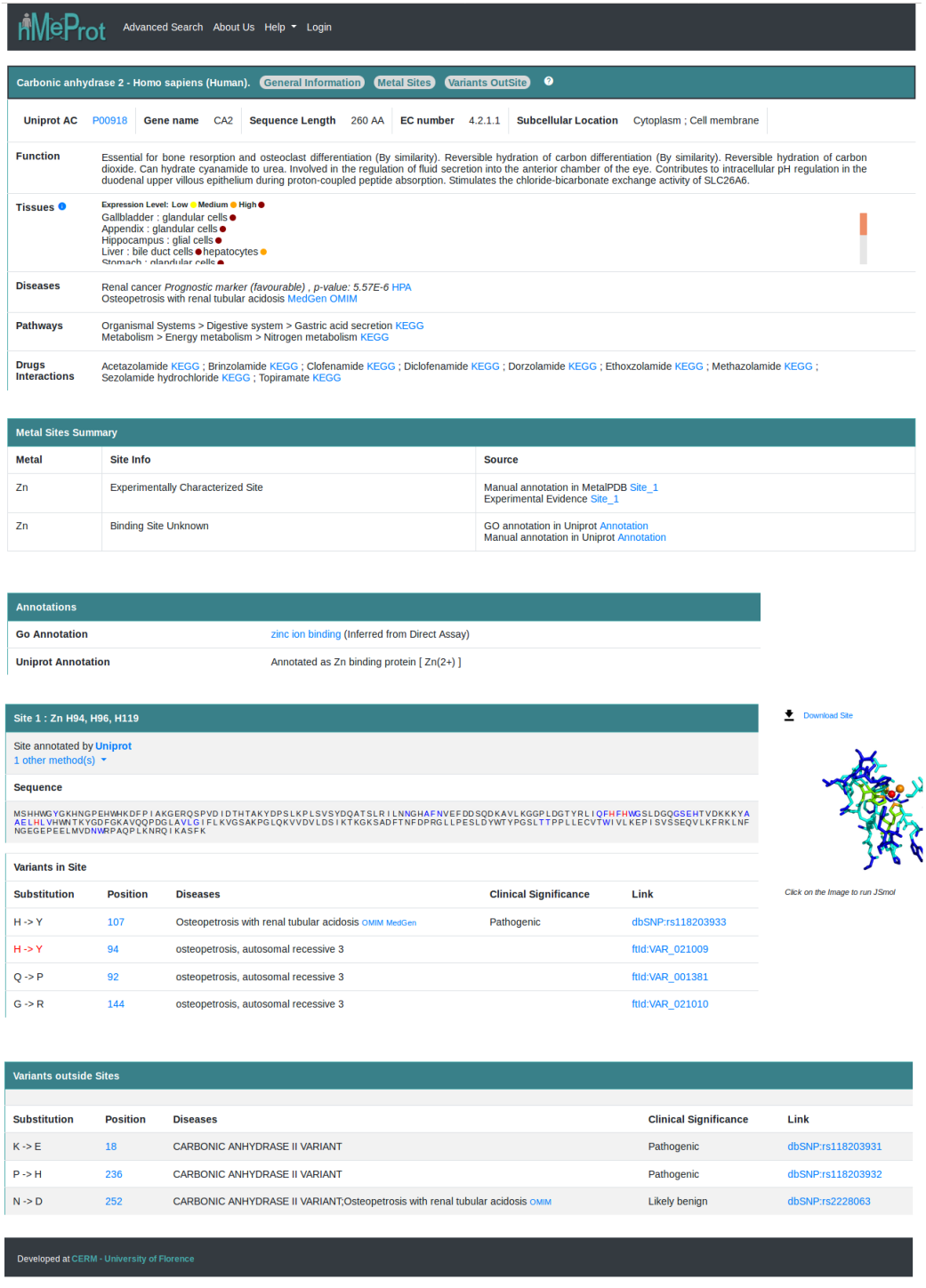
Below the general information, the second section (metal sites) reports a summary of the metal-binding sites found in the protein, grouped by metal type. For each metal-binding site, hMeProt provides information on both the method (site info column) and the type of evidence (source column) used to identify the site. Specifically, the site info column contains one of the following:

- Experimentally characterized site: it defines a site determined by evidence based on experimental data; identification of the site occurred by the literature-based Uniprot method, the MetalPDB method or the structure-based Uniprot method (see Methods).
- Predicted site: it defines a site predicted by bioinformatics analysis; identification of the site occurred by the similarity-based Uniprot method or the Prosite method (see Methods).
- Binding site unknown: it defines a putative site in a protein that was annotated as metal-binding, but lacking information on the metal ligands; identification of the site occurred by the Uniprot without ligands method or the GO method (see Methods).

Each site is associated with a unique identifier in the database, and is linked to a number of site details shown below in the page. These details include: (i) the metal type; (ii) the metal ligands; (iii) the method(s) by which the site was identified; (iv) the complete protein sequence with the metal-binding residues and the neighboring residues stained in red and blue, respectively; and (v) the image of the 3D structure of the site. By clicking on this latter image the user can access an interactive JSmol (wiki.jmol.org/index.php/JSmol) viewer for visualization and examination of the structure of the site. It is also possible to download a PDB file with this structure. Instead, by clicking on the method used to identify the site, the user can obtain more details on how the method was applied: for example, for sites predicted by Prosite is shown the alignment of the protein sequence to the Prosite profile(s), while for sites predicted by MetalPDB are shown the metal ligands in the structures corresponding to the protein. If there are known sequence variants affecting the amino acid residues that form the site, these are also listed in a table displayed below the above details. The table shows, for each sequence variant, (i) the amino acid substitution, (ii) the position on the sequence, (iii) the disease(s) associated with the variant, (iv) the clinical significance of the variant (if available), and (v) the link to an external database (SwissVar[76] or dbSNP[77]). If a variant occurs in the first sphere (i.e., affects one of the ligands) of the site, the amino acid substitution in the first column of the table is displayed in red to make the information more

evident. Furthermore, by clicking on the position of a variant in the sequence (second column of the table), the corresponding residue will be highlighted within the protein sequence above. At the bottom of the page, finally, the third section (sequence variants outside the metal site) reports a table describing the sequence variants that occur outside of all the metal-binding sites, completely analogous to the above described table referred to the variants within the metal site.

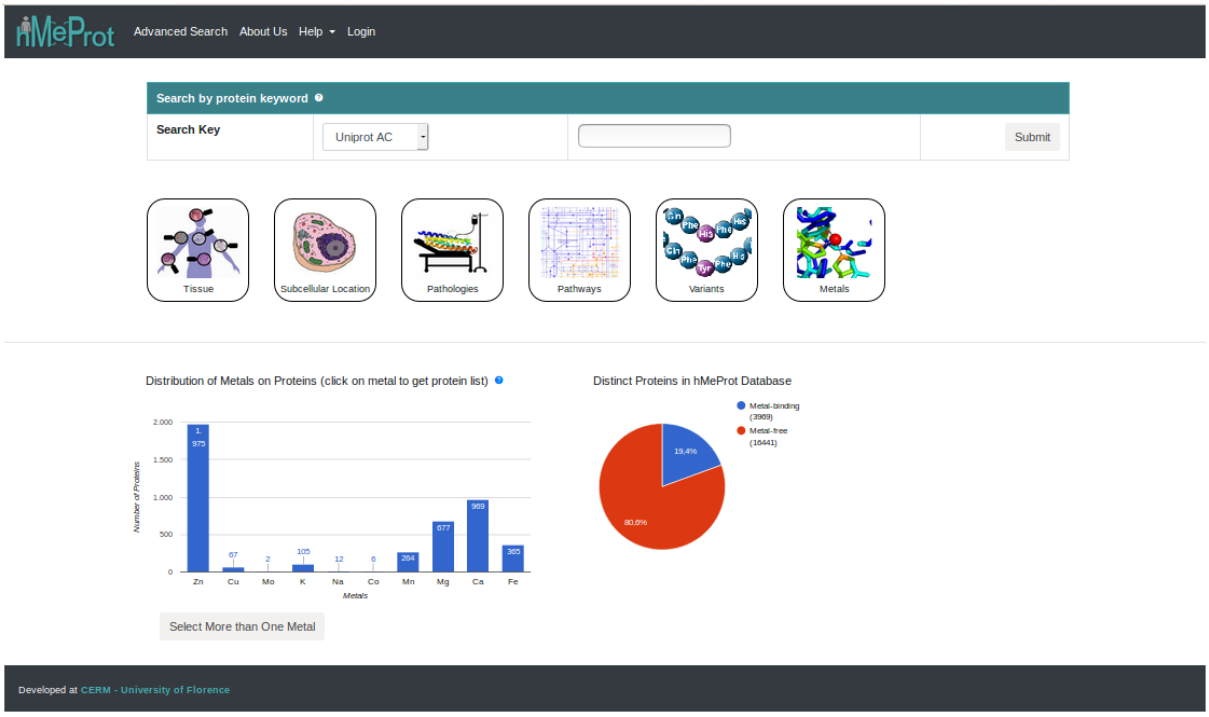
Figure 18. *HMeProt protein page for Carbonic anhydrase 2 (P00918).*



5.3.3 hMeProt statistics pages

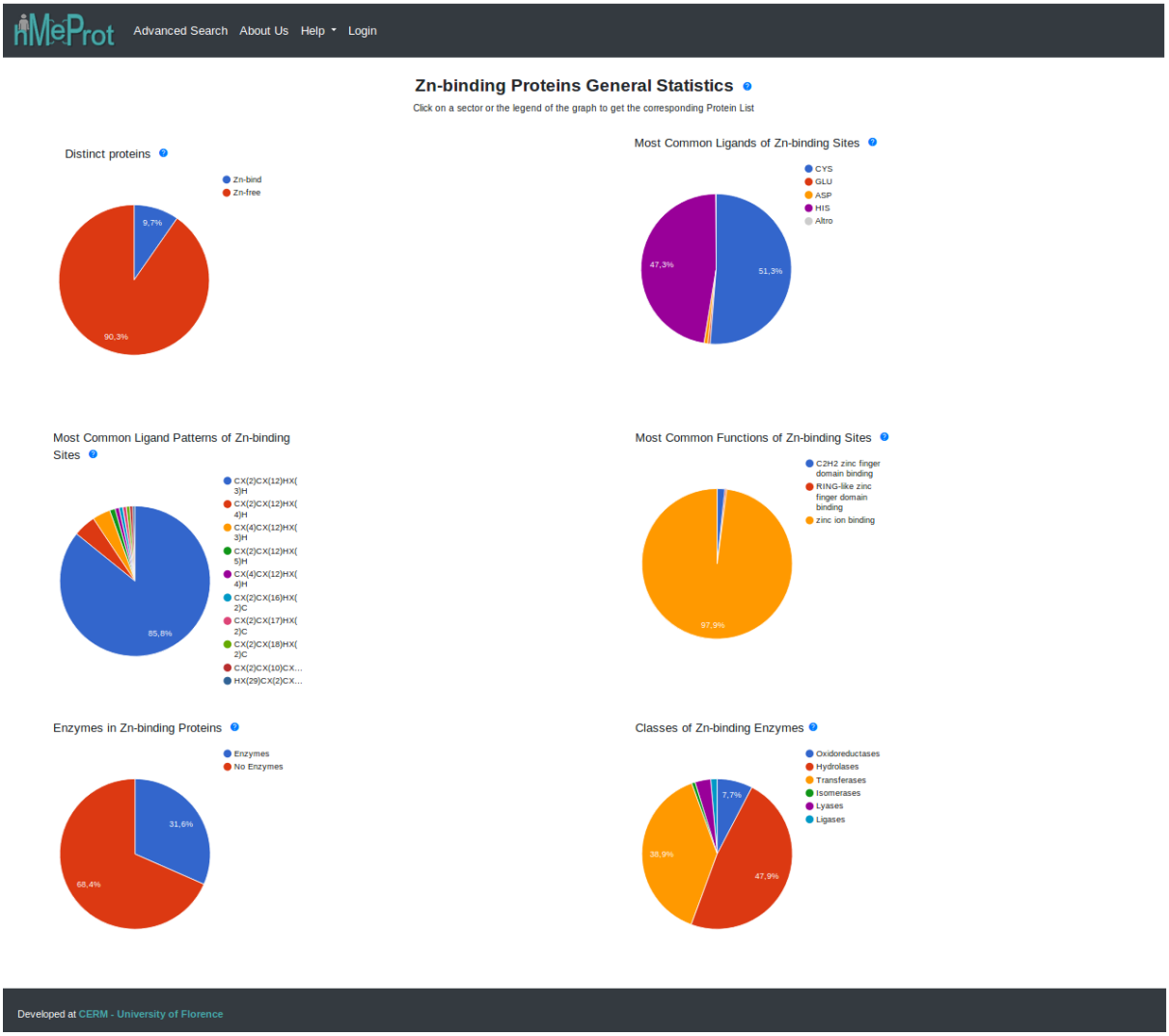
Besides the detailed data provided in the protein pages, users can also access the information on the metal-binding proteins contained in the hMeProt database in an aggregate manner, by means of the statistics pages. These pages are reachable through the six tabs displayed in the home page (Figure 19), which correspond to the criteria by which statistics were built, i.e.(i) tissue(s) in which proteins are expressed, (ii) subcellular location(s) of proteins, (iii) pathologies and (iv) pathways in which proteins are involved, (v) variants in protein sequences, and (vi) type(s) of metal bound.

Figure 19. *Home page of the hMeProt database*



Three of the above tabs, i.e., Tissue, Variants and Metals, allow the user to select a specific metal to obtain a database analysis based on it. For example, by clicking on the Metals tab and then selecting zinc as the metal of interest, the user will obtain a statistics page about zinc-binding proteins and their metal sites (Figure 20). In particular, for iron-binding proteins it is also possible to narrow the analysis to a specific iron cofactor (iron ion, heme iron, or iron sulfur clusters), thereby obtaining separate plots for each case.

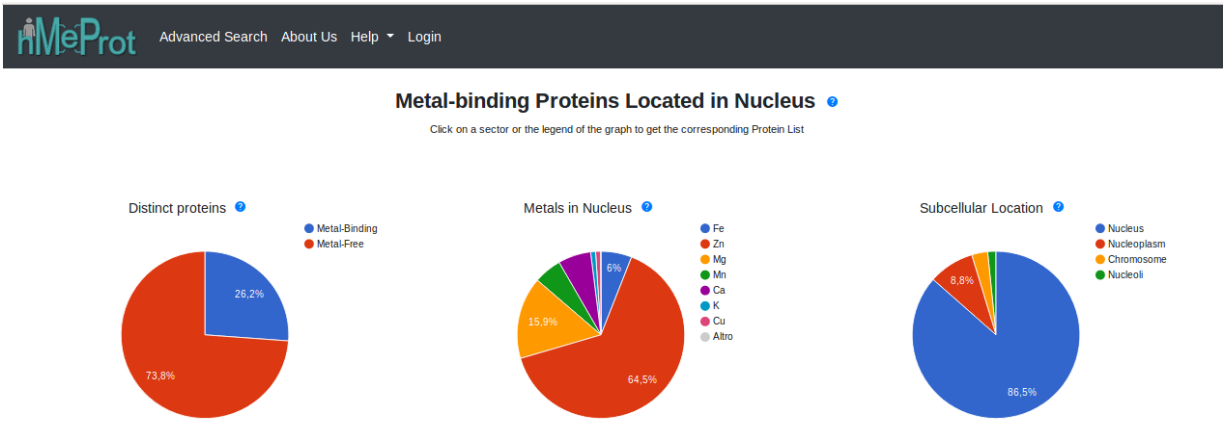
Figure 20. *Statistics page for zinc-binding proteins in the hMeProt database*



The other three tabs, i.e., Subcellular location, Pathologies and Pathways allow, as an alternative to the selection of a specific metal, also the selection of the object of interest (i.e., a specific subcellular location, pathology or pathway, respectively) to obtain statistics in relation to it. For example, by clicking on the Subcellular location tab and then selecting nucleus as the location of interest, the user will obtain a statistics page about the metal-binding proteins found in the nucleus (Figure 21).

The statistics pages in hMeProt present cumulative data regarding not only metal-binding proteins, but also human proteins that are not metal-binding. As previously noted, this is aimed at facilitating users to perform comparative analyses between metal-binding and non-metal-binding proteins.

Figure 21. *Statistics page for nuclear metal binding proteins in the hMeProt database*



5.3.4 Querying the hMeProt database

The web interface of hMeProt offers many options to interrogate the database. From the home page (see Figure 19) it is possible to perform a search in any of the following ways:

(i) by key, i.e., by UniProt[44] accession code, gene name or protein name;

(ii) by one of the six tabs, i.e., by tissue, subcellular location, pathologies, pathways, variants or metals (see section 5.3.3 above);

(iii) by metal type, clicking on a column of the "Distribution of metals on proteins" chart or on the "Select more than one metal" button.

The search by protein name is provided with an autocomplete search system: by typing at least three characters, a list will be shown with the names of all the proteins in hMeProt that match the typed characters.

The results of a search (except the search by key, which leads directly to a protein page) are shown in a metal-binding protein list page (see, e.g., Figure 23). It is possible to filter the results using the input text boxes located at the top of each column, customize the data included in the list (using the "Customize columns" button), and download the results in a csv-formatted file (using the "Download data" button). By clicking on the "Show" button in the last column, the user can select a specific protein and be redirected to the corresponding protein page (see Figure 18 and section 5.3.2 above).

An additional search system in hMeProt is provided by the charts found, for example, in the statistics pages, because all hMeProt charts are interactive. For example, after building a statistics page for zinc-binding proteins based on sequence variants (see section 5.3.3 above),

it is possible to obtain the list of all zinc-binding proteins that have at least one sequence variant in the first sphere of their metal sites simply by clicking on the corresponding sector of the chart (Figure 22 and Figure 23).

Figure 22. *Statistics page for zinc-binding proteins with sequence variants in the hMeProt database. By clicking on the blue sector of the third pie chart from the left, the list shown in Figure 23 will be generated.*

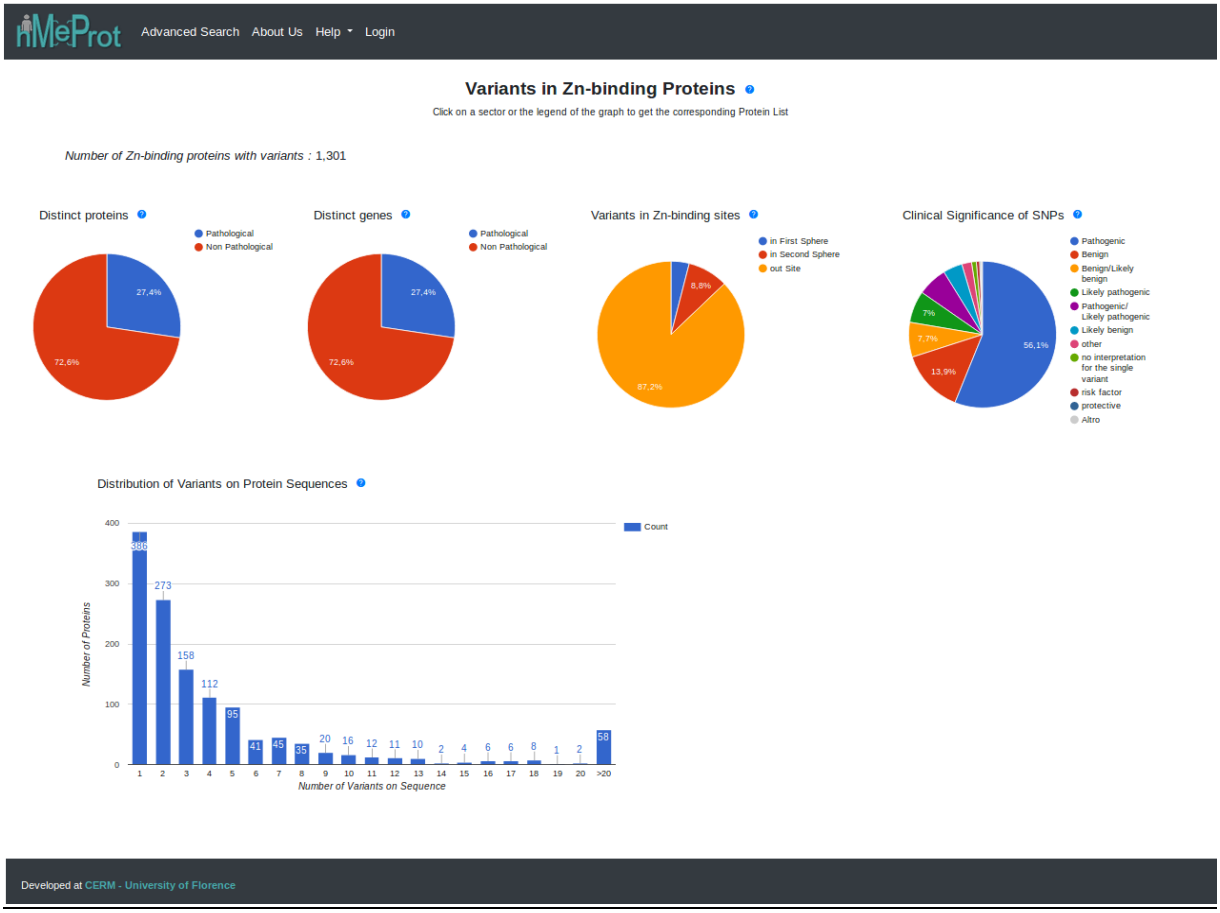
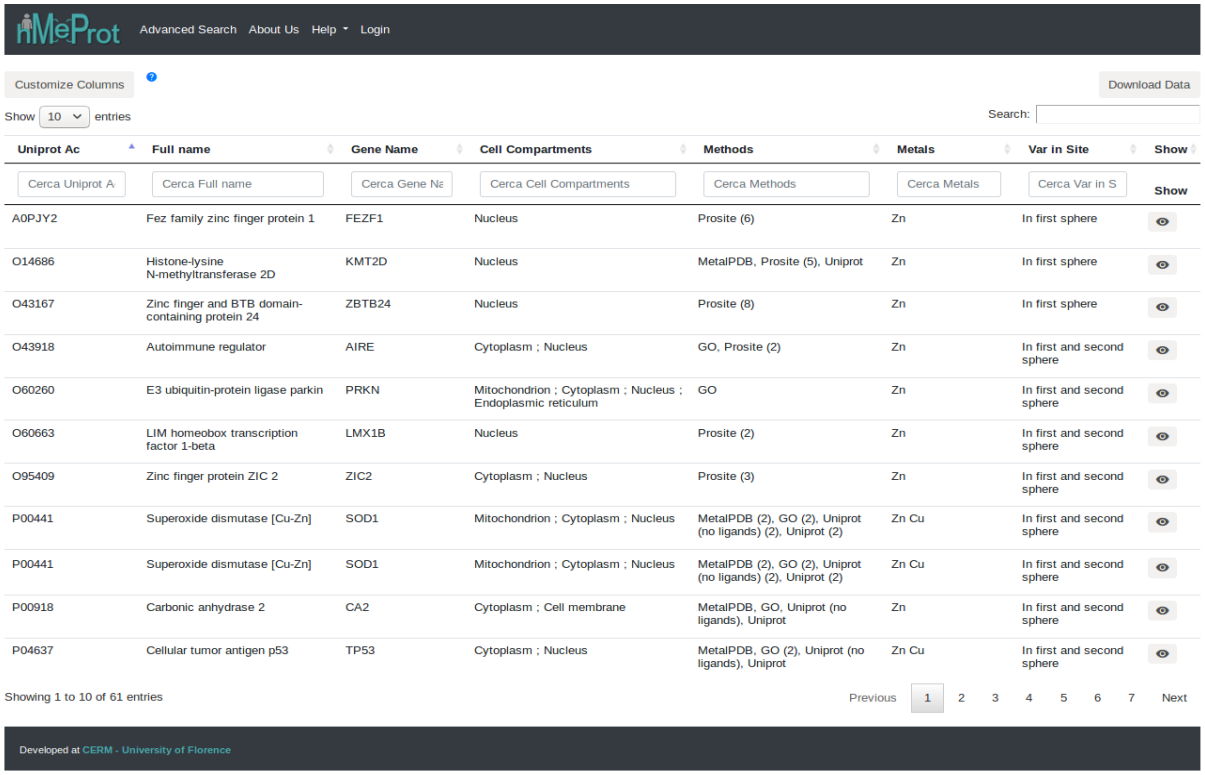
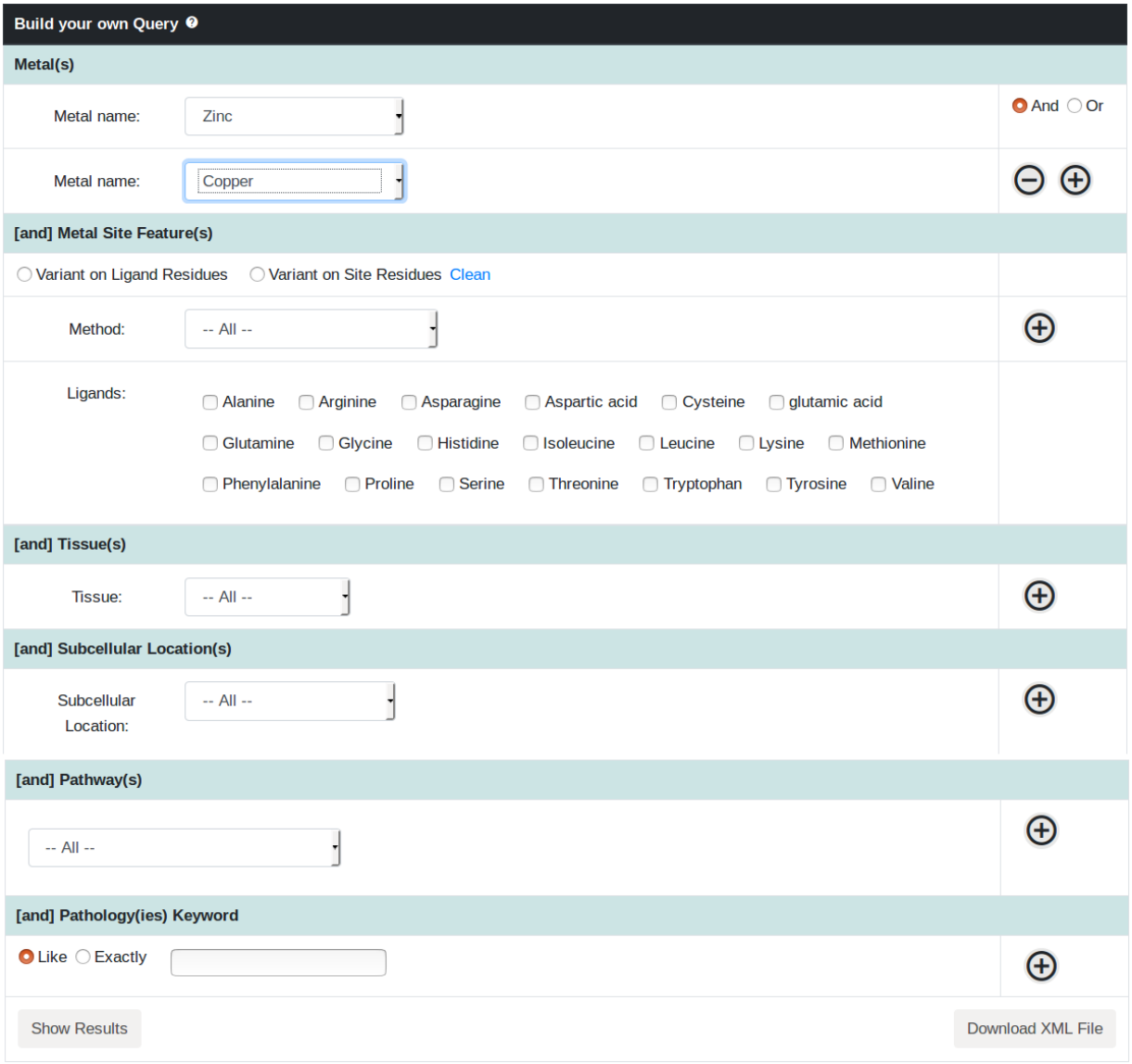


Figure 23. *List of zinc-binding proteins generated as described in the legend of Figure 22.*



Finally, the web interface of hMeProt provides an advanced search option to perform custom queries to the database (Figure 24), thereby allowing users to search for terms in one or more specific fields of choice. The Query Builder of hMeProt is organized by sections (Metals, Metal Site Features, Tissues, Subcellular locations, Pathways, Pathologies), and it is possible to use more than one search key in the same section. The Pathologies section is provided with a "like" option operator, to allow the user to search proteins associated with pathologies whose names match the given pattern. The Pathways section has a guided search: after the selection of a pathway, the list of all the sub-pathways related to it will be shown, and the same will happen upon the selection of a sub-pathway, as long as a further sub-level of pathways exists. The results of an advanced search can also be downloaded as an XML file, and the user can select the information to be included in the report. The data available for selection can refer both to proteins (e.g., protein name, gene name, function) and to metal sites (e.g., metal bound, pattern of metal ligands).

Figure 24. *Example of advanced search in hMeProt with two metals selected for search. The boolean operator allowed between different sections is only "AND", while between fields of the same section both "AND" and "OR" are allowed.*



5.3.5 Final considerations on the hMeProt database

The hMeProt database is a new resource I developed to provide an overview of the human metalloproteome. It collects all the human metal-binding proteins identified by experimental or bioinformatics methods. The latter represent a substantial contribution to the definition of the metalloproteome, given the difficulty and cost of performing empirical investigations on metal-binding sites at the whole proteome scale [89]. Furthermore, metal sites could be defined for the majority of the predicted metal-binding proteins, and the level of

detail arrives at the identification in the sequence of the metal ligand residues and their neighboring residues. This makes possible, starting from the experimental structure of the protein or a 3D template, to reconstruct the 3D structure of predicted sites. Using the web interface of hMeProt, the structures of metal sites can be explored interactively, and are available for download.

In hMeProt the human proteins are framed in the organismal and cellular context, and are connected both with the biological pathways and with the diseases in which they are involved. This kind of information will be useful, for example, to understand the cellular processes affected by the deficiency or the dysregulation of given metal ions, and thus the consequences for the organism. In this regard, it is important to note that hMeProt collects information about all human proteins, not only about the metal-binding proteins. This allows users to focus on metal-containing players in cellular processes, yet avoiding to narrow down their analysis to them alone.

A key feature of hMeProt is that it allows one to examine the relationship between sequence variants (especially SNPs) associated with human disease and metal-binding sites in proteins [90]. By integrating the data concerning the variants present in the protein sequences with the sequence positions of the residues forming metal sites, it makes possible to study the effect of amino acid substitutions on the interaction with the metal, as well as, by further providing information on the pathologies associated with each variant, the possible roles of impaired metal sites in human diseases.

Finally, the large amount of statistical analyses provided on the resource and the many ways to query the database make hMeProt a very versatile tool for the study of the human metalloproteome. Thanks to the combination of different expertise across bioinorganic chemistry, bioinformatics, statistics, and computational chemistry, hMeProt will provide the scientific community with an unprecedented information on the human metalloproteome, thus contributing to shed light on the roles of metal ions in healthy metabolism and under pathological conditions, and supporting the growing needs of bioinorganic chemists to store, manage, share and process proteomics data.

6 CONCLUSIONS

Metal-binding proteins, i.e. proteins that bind a metal ion to carry out their physiological function, are essential to life. Current data indicate that about 40% of structures in the PDB are metal-binding proteins, and about 40% of enzymes with known structure use a metal ion to carry out the reaction mechanism [42]. In fact, metal-binding proteins participate to the most important biochemical processes, including respiration, nitrogen fixation and photosynthesis [86,91,92].

For a long period, bioinformatics has almost completely neglected to develop resources and tools to study the interaction between metal ions and proteins, probably because metal sites are difficult to encode with models. Currently, the most exhaustive available resource focused on metals in biology is MetalPDB, a database on which I have worked during my Ph.D. This resource, based on the concept of the Minimal Functional Site (MFS), *is* aimed at providing the scientific community with all the available information on metal sites in protein structures. MetalPDB provides an exhaustive overview of the roles of metals in proteins, exploring the sequence-structure and structure-function relationships in MFSs. The thoroughness of MetalPDB has made it one of the reference resources for the study of metals in biology. The growth over the years in the interest by the scientific community is revealed by the increase in the contacts to the database (Figure 13), which in 2018 reached an average of almost 4000 visits each month.

MetalPDB also acts as a platform where users have free access to a number of tools designed to study metal-binding proteins using MFSs as the central concept. One of these is MetalPredator, a web server to predict iron-, zinc- and copper-binding sites in protein sequences on which I worked during my Ph.D., too. This tool integrates global and local searches to recognize metal sites in sequences, using an approach that overcomes most of the limitations of the current methods for the prediction of metalloproteomes, and thus has a higher coverage. The major strengths of the MetalPredator approach are that it is based on flexible rather than rigid metal-binding patterns, therefore it has the potential to also predict metal-mediated protein-protein interactions, metal-sites in IDPs and regulatory sites. Using MetalPredator, we were thus able to predict the human iron-proteome with high accuracy.

The challenge in the study of metalloproteomes is not only the identification of metal-binding proteins, but also the understanding of how metal ions and metal-binding molecules, together with all other cellular components, contribute to the metabolism of healthy cells and, under pathological conditions, lead to the onset of metal-associated diseases. The study of the

human metalloproteome is especially relevant to this task, therefore during my Ph.D. I have also worked on the development of the hMeProt database. hMeProt is a resource that integrates the human metalloproteome data with various other types of information, so as to frame each metal-binding protein into the cellular/pathological context. In addition, the high level of detail at which metal sites in human proteins are defined in the database makes hMeProt an ideal resource to investigate both the structure-function relationships in metal-binding proteins and the influence of genetic variations on metal site properties.

In conclusion, we expect that the resources developed within this doctorate will provide a valuable support to a wide range of scientists involved in the study of metals in biology; besides producing novel data on metalloproteomes, they will facilitate the access to integrated data, assisting the process of knowledge discovery and ultimately enhancing our understanding of the fascinating, inextricable link between life and the inorganic world of metal ions.

Reference List

1 Bertini, I., Sigel, A. & Sigel, H. *Handbook on Metalloproteins*. 1 edn, Vol. 1 1-1800 (Marcel Dekker, 2001).

2 Mertz, W. The essential trace elements. *Science* **213**, 1332-1338, doi:10.1126/science.7022654 (1981).

3 Mills, C. F. Trace elements in animals. *Philos Trans R Soc Lond B Biol Sci* **288**, 51-63, doi:10.1098/rstb.1979.0090 (1979).

4 Rincker, M. J., Hill, G. M., Link, J. E., Meyer, A. M. & Rowntree, J. E. Effects of dietary zinc and iron supplementation on mineral excretion, body composition, and mineral status of nursery pigs. *J Anim Sci* **83**, 2762-2774, doi:10.2527/2005.83122762x (2005).

5 Scheplyagina, L. A. Impact of the mother's zinc deficiency on the woman's and newborn's health status. *J Trace Elem Med Biol* **19**, 29-35, doi:10.1016/j.jtemb.2005.07.008 (2005).

6 Kilicalp, D., Dede, S., Belge, F. & Tatar, M. Effect of protein deficiency on macroelement and trace element levels of weanling rats' small intestine and liver tissues. *Biol Trace Elem Res* **107**, 255-261, doi:10.1385/BTER:107:3:255 (2005).

7 Morgan, J. Nutrition for toddlers: the foundation for good health--1. toddlers' nutritional needs: what are they and are they being met? *J Fam Health Care* **15**, 56-59 (2005).

8 Puntarulo, S. Iron, oxidative stress and human health. *Mol Aspects Med* **26**, 299-312, doi:10.1016/j.mam.2005.07.001 (2005).

9 Nielsen, F. H. Evolutionary events culminating in specific minerals becoming essential for life. *Eur. J. Nutr* **39**, 62-66 (2000).

10 Arredondo, M. & Nunez, M. T. Iron and copper metabolism. *Mol Aspects Med* **26**, 313-327, doi:10.1016/j.mam.2005.07.010 (2005).

11 Finney, L. A. & O'Halloran, T. V. Transition Metal Speciation in the Cell: Insights from the Chemistry of Metal Ion Receptors. *Science* **300**, 931-936 (2003).

12 Outten, C. E. & O'Halloran, T. V. Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science* **292**, 2488-2492 (2001).

13 Andrews, S. C., Robinson, A. K. & Rodriguez-Quiñones, F. Bacterial iron homeostasis. *FEMS Microbiol Rev* **27**, 215-237 (2003).

14 Silver, S. & Phung le, T. A bacterial view of the periodic table: genes and proteins for toxic inorganic ions. *J Ind Microbiol Biotechnol* **32**, 587-605, doi:10.1007/s10295-005-0019-6 (2005).

15 Tottey, S., Harvie, D. R. & Robinson, N. J. Understanding how cells allocate metals using metal sensors and metallochaperones. *Accounts of Chemical Research* **38**, 775-783 (2005).

16 Torti, S. V. & Torti, F. M. Cellular iron metabolism in prognosis and therapy of breast cancer. *Crit Rev. Oncog* **18**, 435-448 (2013).

17 Malemud, C. J. Matrix metalloproteinases (MMPs) in health and disease: an overview. *Front Biosci* **11:1696-701.**, 1696-1701 (2006).

18 Perez-Llarena, F. J. & Bou, G. Beta-lactamase inhibitors: the story so far. *Curr. Med. Chem* **16**, 3740-3765 (2009).

19 Holm, R. H., Kennepohl, P. & Solomon, E. I. Structural and Functional Aspects of Metal Sites in Biology. *Chem Rev* **96**, 2239-2314 (1996).

20 Maret, W. Zinc coordination environments in proteins determine zinc functions. *J Trace Elem Med Biol* **19**, 7-12, doi:10.1016/j.jtemb.2005.02.003 (2005).

21 Higgins, C. L., Muralidhara, B. K. & Wittung-Stafshede, P. How do cofactors modulate protein folding? *Protein Pept Lett* **12**, 165-170 (2005).

22 Sieker, L. C. *et al.* Alternative metal-binding sites in rubrerythrin. *Nat Struct Biol* **6**, 308-309, doi:10.1038/7538 (1999).

23 Cotton, F. A. & Wilkinson, G. *Advanced Inorganic Chemistry*. 4 edn, (Wiley, 1990).

24 Degtyarenko, K. N., North, A. C. T. & Findlay, J. B. C. PROMISE: a database of bioinorganic motifs. *Nucleic Acids Res* **27**, 233-236 (1999).

25 Castagnetto, J. M. *et al.* MDB: the Metalloprotein Database and Browser at The Scripps Research Institute. *Nucleic Acids Res* **30**, 379-382 (2002).

26 Degtyarenko, K. & Contrino, S. COMe: the ontology of bioinorganic proteins. *BMC. Struct. Biol* **4**, 3 (2004).

27 Hsin, K., Sheng, Y., Harding, M. M., Taylor, P. & Walkinshaw, M. D. MESPEUS: a database of the geometry of metal sites in proteins. *J. Appl. Cryst* **41**, 963-968 (2008).

28 Rose, P. W. *et al.* The RCSB protein data bank: integrative view of protein, gene and 3D structural information. *Nucleic Acids Res* **45**, D271-D281, doi:gkw1000 [pii];10.1093/nar/gkw1000 [doi] (2017).

29 Choi, H., Kang, H. & Park, H. MetLigDB: a web-based database for identification of chemical groups to design metalloprotein inhibitors. *J. Appl. Cryst* **44**, 878-881 (2011).

30 Schnabl, J., Suter, P. & Sigel, R. K. O. MINAS--a database of Metal Ions in Nucleic AcidS. *Nucleic Acids Res* **40**, D434-D438 (2012).

31 Andreini, C., Cavallaro, G., Lorenzini, S. & Rosato, A. MetalPDB: a database of metal sites in biological macromolecular structures. *Nucleic Acids Res* **41**, D312-D319 (2013).

32 Andreini, C., Cavallaro, G. & Lorenzini, S. FindGeo: a tool for determining metal coordination geometry. *Bioinformatics* **28**, 1658-1660 (2012).

33 Bertini, I., Luchinat, C., Provenzani, A., Rosato, A. & Vasos, P. R. Browsing gene banks for Fe_2S_2 ferredoxins and structural modeling of 87 plant-type sequences: an analysis of fold and function. *Proteins: Structure, Function, and Genetics* **46**, 110-127 (2002).

34 Schmidt, M., Meier, B. & Parak, F. X-ray Structure of the Cambialistic Superoxide Dismutase from Propionibacterium shermanii Active with Fe or Mn. *JBIC* **1**, 532-541 (1996).

35 Waldron, K. J. & Robinson, N. J. How do bacterial cells ensure that metalloproteins get the correct metal? *Nat. Rev. Microbiol* **7**, 25-35 (2009).

36 Banci, L. & Piccioli, M. in *Encyclopedia of Magnetic Resonance* (eds D.M. Grant & R.K. Harris) 1365-1373 (Encycl. of Nuclear Magn. Reson., 1996).

37 D'Onofrio, M. *et al.* High relaxivity supramolecular adducts between human-liver fatty-acid-binding protein and amphiphilic Gd(III) complexes: structural basis for the design of intracellular targeting MRI probes. *Chemistry* **18**, 9919-9928 (2012).

38 Barrick, D. Depletion and replacement of protein metal ligands. *Curr. Opin. Biotechnol* **6**, 411-418 (1995).

39 Lu, Y., Yeung, N., Sieracki, N. & Marshall, N. M. Design of functional metalloproteins. *Nature* **460**, 855-862 (2009).

40 Swingle, M. R., Honkanen, R. E. & Ciszak, E. M. Structural basis for the catalytic activity of human serine/threonine protein phosphatase-5. *J. Biol. Chem* **279**, 33992-33999 (2004).

41 Shi, Y. Serine/threonine phosphatases: mechanism through structure. *Cell* **139**, 468-484 (2009).

42 Andreini, C., Bertini, I., Cavallaro, G., Holliday, G. L. & Thornton, J. M. Metal ions in biological catalysis: from enzyme databases to general principles. *J. Biol. Inorg. Chem* **13**, 1205-1218 (2008).

43 McDonald, I. K. & Thornton, J. M. Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol* **238**, 777-793 (1994).

44 Consortium, U. The Universal Protein Resource (UniProt) 2009. *Nucleic Acids Res* **37**, D169-D174 (2009).

45 Cuff, A. L. *et al.* Extending CATH: increasing coverage of the protein structure universe and linking structure with function. *Nucleic Acids Res* **39**, D420-D426 (2011).

46 Andreeva, A. *et al.* Data growth and its impact on the SCOP database: new developments. *Nucleic Acids Res* **36**, D419-D425 (2008).

47 Punta, M. *et al.* The Pfam protein families database. *Nucleic Acids Res* **40**, D290-D301 (2012).

48 Fu, D. & Finney, L. Metalloproteomics: challenges and prospective for clinical research applications. *Expert Rev. Proteomics* **11**, 13-19 (2014).

49 Tomalova, I., Foltynova, P., Kanicky, V. & Preisler, J. MALDI MS and ICP MS detection of a single CE separation record: a tool for metalloproteomics. *Anal. Chem* **86**, 647-654 (2014).

50 Shi, W. & Chance, M. R. Metalloproteomics: forward and reverse approaches in metalloprotein structural and functional characterization. *Curr. Opin. Chem. Biol* **15**, 144-148 (2011).

51 Shu, N., Zhou, T. & Hovmoller, S. Prediction of zinc-binding sites in proteins from sequence. *Bioinformatics* **24**, 775-782 (2008).

52 Dupont, C. L., Yang, S., Palenik, B. & Bourne, P. E. Modern proteomes contain putative imprints of ancient shifts in trace metal geochemistry. *Proc. Natl. Acad. Sci. USA* **103**, 17822-17827 (2006).

53 Zhang, Y. & Gladyshev, V. N. Molybdoproteomes and evolution of molybdenum utilization. *J. Mol. Biol* **379**, 881-899 (2008).

54 Passerini, A., Lippi, M. & Frasconi, P. MetalDetector v2.0: predicting the geometry of metal binding sites from protein sequence. *Nucleic Acids Res* (2011).

55 Andreini, C., Bertini, I. & Rosato, A. Metalloproteomes: a bioinformatic approach. *Accounts of Chemical Research* **42**, 1471-1479 (2009).

56 Chen, Z., Wang, Y., Zhai, Y. F., Song, J. & Zhang, Z. ZincExplorer: an accurate hybrid method to improve the prediction of zinc-binding sites from protein sequences. *Mol. Biosyst* **9**, 2213-2222 (2013).

57 Zhang, Y., Rump, S. & Gladyshev, V. N. Comparative Genomics and Evolution of Molybdenum Utilization. *Coord. Chem. Rev* **255**, 1206-1217 (2011).

58 Shi, W. *et al.* Characterization of metalloproteins by high-throughput X-ray absorption spectroscopy. *Genome Res* **21**, 898-907 (2011).

59 Gladyshev, V. N. & Zhang, Y. Comparative genomics analysis of the metallomes. *Met. Ions. Life Sci* **12:529-80. doi: 10.1007/978-94-007-5561-1_16.**, 529-580 (2013).

60 Andreini, C., Bertini, I. & Rosato, A. A hint to search for metalloproteins in gene banks. *Bioinformatics* **20**, 1373-1380 (2004).

61 Passerini, A., Lippi, M. & Frasconi, P. Predicting metal-binding sites from protein sequence. *IEEE/ACM Trans Comput Biol Bioinform* **9**, 203-213, doi:10.1109/TCBB.2011.94 (2012).

62 Valasatava, Y., Rosato, A., Banci, L. & Andreini, C. Metalpredator: a web server to predict iron-sulfur cluster binding proteomes. *Bioinformatics* **btw 238** (2016).

63 Hutchinson, E. G. & Thornton, J. M. PROMOTIF--a program to identify and analyze structural motifs in proteins. *Protein Sci* **5**, 212-220, doi:10.1002/pro.5560050204 [doi] (1996).

64 Notredame, C., Higgins, D. G. & Heringa, J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol* **302**, 205-217 (2000).

65 Andreini, C., Cavallaro, G., Rosato, A. & Valasatava, Y. MetalS^2: a tool for the structural alignment of minimal functional sites in metal-binding proteins and nucleic acids. *J. Chem. Inf. Model* **53**, 3064-3075 (2013).

66 Andreini, C., Bertini, I., Cavallaro, G., Decaria, L. & Rosato, A. A simple protocol for the comparative analysis of the structure and occurence of biochemical pathways across superkingdoms. *J. Chem. Inf. Model* **51**, 730-738 (2011).

67 Schaffer, A. A. *et al.* Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements. *Nucleic Acids Res* **29**, 2994-3005, doi:10.1093/nar/29.14.2994 (2001).

68 Suzek, B. E., Huang, H., McGarvey, P., Mazumder, R. & Wu, C. H. UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics* **23**, 1282-1288, doi:10.1093/bioinformatics/btm098 (2007).

69 Eddy, S. R. Profile hidden Markov models. *Bioinformatics* **14**, 755-763 (1998).

70 Johnson, L. S., Eddy, S. R. & Portugaly, E. Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinformatics* **11**, 431, doi:10.1186/1471-2105-11-431 (2010).

71 Huang, Y., Niu, B., Gao, Y., Fu, L. & Li, W. CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* **26**, 680-682, doi:10.1093/bioinformatics/btq003 (2010).

72 Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402 (1997).

73 Hulo, N. *et al.* The PROSITE database. *Nucleic Acids Res* **34**, D227-230, doi:10.1093/nar/gkj063 (2006).

74 The Gene Ontology project in 2008. *Nucleic Acids Res* **36**, D440-D444 (2008).

75 Uhlen, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419, doi:10.1126/science.1260419 (2015).

76 Yip, Y. L. *et al.* Annotating single amino acid polymorphisms in the UniProt/Swiss-Prot knowledgebase. *Hum Mutat* **29**, 361-366, doi:10.1002/humu.20671 (2008).

77 Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* **29**, 308-311, doi:10.1093/nar/29.1.308 (2001).

78 Landrum, M. J. & Kattman, B. L. ClinVar at five years: Delivering on the promise. *Hum Mutat* **39**, 1623-1630, doi:10.1002/humu.23641 (2018).

79 Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**, D353-D361, doi:10.1093/nar/gkw1092 (2017).

80 Maglott, D., Ostell, J., Pruitt, K. D. & Tatusova, T. Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Res* **39**, D52-57, doi:10.1093/nar/gkq1237 (2011).

81 Amberger, J. S. & Hamosh, A. Searching Online Mendelian Inheritance in Man (OMIM): A Knowledgebase of Human Genes and Genetic Phenotypes. *Curr Protoc Bioinformatics* **58**, 1 2 1-1 2 12, doi:10.1002/cpbi.27 (2017).

82 Sayers, E. W. *et al.* Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **47**, D23-D28, doi:10.1093/nar/gky1069 (2019).

83 Weinreich, S. S., Mangon, R., Sikkens, J. J., Teeuw, M. E. & Cornel, M. C. [Orphanet: a European database for rare diseases]. *Ned Tijdschr Geneeskd* **152**, 518-519 (2008).

84 Haas, J. *et al.* The Protein Model Portal--a comprehensive resource for protein structure and model information. *Database (Oxford)* **2013**, bat031, doi:10.1093/database/bat031 (2013).

85 Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296-W303, doi:10.1093/nar/gky427 (2018).

86 Yruela, I. Transition metals in plant photosynthesis. *Metallomics* **5**, 1090-1109, doi:10.1039/c3mt00086a (2013).

87 Andreini, C., Bertini, I. & Cavallaro, G. Minimal functional sites allow a classification of zinc sites in proteins. *Plos ONE* **10**, e26325 (2011).

88 Sillitoe, I. *et al.* CATH: comprehensive structural and functional annotations for genome sequences. *Nucleic Acids Res* **43**, D376-D381, doi:gku947 [pii];10.1093/nar/gku947 [doi] (2015).

89 Barnett, J. P., Scanlan, D. J. & Blindauer, C. A. Protein fractionation and detection for metalloproteomics: challenges and approaches. *Anal Bioanal Chem* **402**, 3311-3322, doi:10.1007/s00216-012-5743-y (2012).

90 Levy, R., Sobolev, V. & Edelman, M. First- and second-shell metal binding residues in human proteins are disproportionately associated with disease-related SNPs. *Hum Mutat* **32**, 1309-1318, doi:10.1002/humu.21573 (2011).

91 Frey, P. A. Advances in research on metalloproteins. *Methods Mol Biol* **1122**, 1-3, doi:10.1007/978-1-62703-794-5_1 (2014).

92 David, S. S. & Meggers, E. Inorganic chemical biology: from small metal complexes in biological systems to metalloproteins. *Curr Opin Chem Biol* **12**, 194-196, doi:10.1016/j.cbpa.2008.03.008 (2008).