Co-occurrence Analysis

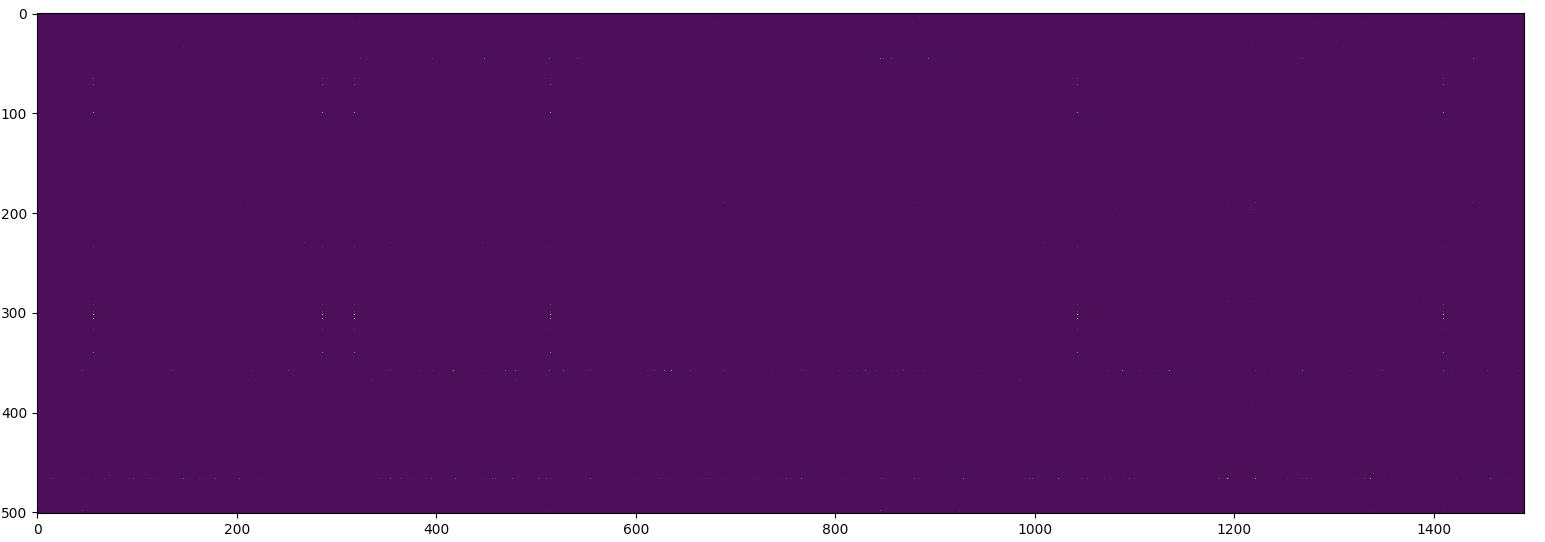


Figure 1: Rows are genes columns are HPO terms. The brighter the more co-occurrences. This is the full heat map, (very small pixels that probably look like dust on your screen are peaks).

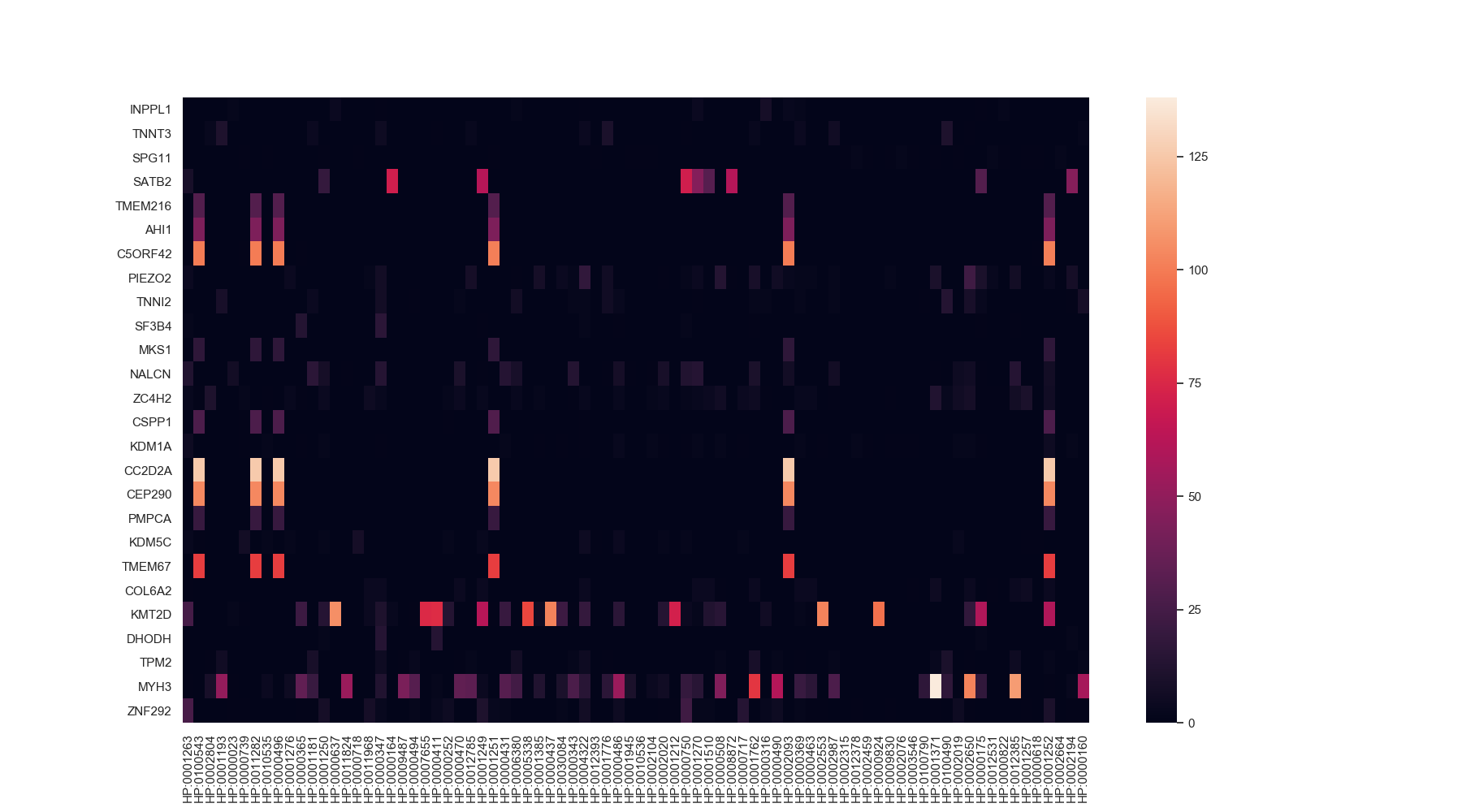


Figure 2: Filtered Heat Map, rows and columns with sums lower than 100 have been removed - 475 rows and 1411 columns were dropped. Interestingly there are several HPOs with very similar gene co-occurrences. Table 1 has an itemized list of the top 20 peaks.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| MYH3 | embryonic skeletal muscle | HP:0001371 | Flexion contracture | 138 |
| CC2D2A | oiled-coil and calcium binding domain | HP:0001252 | Muscular hypotonia | 126 |
| CC2D2A | “ | HP:0002093 | Respiratory insufficiency | 126 |
| CC2D2A | “ | HP:0001251 | Ataxia | 126 |
| CC2D2A | “ | HP:0000496 | Abnormality of eye movement | 126 |
| CC2D2A | “ | HP:0011282 | Abnormality of hindbrain morphology | 126 |
| CC2D2A | “ | HP:0100543 | Cognitive impairment | 126 |
| MYH3 | embryonic skeletal muscle | HP:0012385 | Camptodactyly | 110 |
| KMT2D | lysine-specific methyltransferase | HP:0000637 | Long palpebral fissure | 106 |
| CEP290 | Unknown abundant protein | HP:0001252 | Muscular hypotonia | 104 |
| CEP290 | “ | HP:0002093 | Respiratory insufficiency | 104 |
| CEP290 | “ | HP:0001251 | Ataxia | 104 |
| CEP290 | “ | HP:0000496 | Abnormality of eye movement | 104 |
| CEP290 | “ | HP:0011282 | Abnormality of hindbrain morphology | 104 |
| CEP290 | “ | HP:0100543 | Cognitive impairment | 104 |
| MYH3 | embryonic skeletal muscle | HP:0002650 | Scoliosis | 102 |
| KMT2D | lysine-specific methyltransferase | HP:0002553 | Highly arched eyebrow | 102 |
| KMT2D | “ | HP:0000437 | Depressed nasal tip | 102 |
| C5ORF42 | lysine-specific methyltransferase | HP:0001252 | Muscular hypotonia | 101 |
| C5ORF42 | “ | HP:0001251 | Ataxia | 101 |

Table 1: Gene HPO Term pairs with highest co-occurrence counts. These correspond with the brightest points in Figure 2.

|  |  |
| --- | --- |
| HP:0001252 | Muscular hypotonia |
| HP:0001251 | Ataxia |
| HP:0100543 | Cognitive impairment |
| HP:0002093 | Respiratory insufficiency |
| HP:0000496 | Abnormality of eye movement |
| HP:0011282 | Abnormality of hindbrain morphology |

Table 2: Most common HPO terms

Tops Genes

|  |  |
| --- | --- |
| MYH3 | embryonic skeletal muscle myosin heavy chain 3 |
| KMT2D | lysine |
| CC2D2A | coiled-coil and calcium binding domain protein that appears to play a critical role in cilia formation |
| CEP290 | Abundant but unknown protien |
| C5ORF42 | transmembrane protein |
| SATB2 | protein that helps control the development of certain body systems |
| TMEM67 | gene localizes to the primary cilium and to the plasma membrane |
| PIEZO2 | large transmembrane proteins |
| ZNF292 | Zinc Finger Protein |
| ZC4H2 | C-terminal zinc finger |

Table 3: Most common Genes

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| id | Gene | HPO | phi | gene\_sum | hpo\_sum | well\_known |
| 9 | SATB2 | HP:0000164 | 0.36282 | 506 | 74 | yes |
| 12 | DHODH | HP:0000625 | 0.352744 | 112 | 14 | yes |
| 13 | SATB2 | HP:0008872 | 0.340669 | 506 | 62 | yes |
| 15 | LRP1 | HP:0000121 | 0.33317 | 12 | 3 | yes |
| 16 | LRP1 | HP:0006530 | 0.33317 | 12 | 3 | yes |
| 17 | TMEM231 | HP:0000817 | 0.33317 | 12 | 3 | yes |
| 18 | GLB1 | HP:0008454 | 0.33317 | 12 | 3 | yes |
| 19 | IFT140 | HP:0001636 | 0.33317 | 12 | 3 | yes |
| 20 | DHODH | HP:0000202 | 0.32973 | 112 | 16 | yes |
| 21 | DHODH | HP:0000448 | 0.32973 | 112 | 16 | yes |
| 22 | SF3B4 | HP:0011800 | 0.322677 | 122 | 20 | yes |
| 23 | SS18L1 | HP:0002190 | 0.308427 | 14 | 3 | no |
| 24 | SS18L1 | HP:0002242 | 0.308427 | 14 | 3 | no |
| 25 | COG5 | HP:0100512 | 0.308427 | 14 | 3 | yes |
| 26 | COG5 | HP:0010522 | 0.308427 | 14 | 3 | yes |

Phi-coefficient – Matthew’s Correlations Co-efficient Analysis

Table 4: Gene – HPO pairs with evidence of some correlation based on the phi-coefficient (phi > 0.30). Pairs with an hpo\_sum less than 3 were excluded from the list. Low frequency of use is HPO terms seems to skew the phi. The columns gene\_sum and hpo\_sum are the number of time that a gene or hpo occurred in MyGene2. The last column (well\_known) is a summary of the information found on <https://ghr.nlm.nih.gov/> about the gene and the HPO term descriptions. If there is a seemingly known- well understood connection between the gene and HPO term or not.

**SATB2 & HP:0000164**

HPO: Abnormality of the dentition: Any abnormality of the primary (deciduous) or permanent teeth.

Gene: Mutations in the *SATB2* gene have been found to cause *SATB2*-associated syndrome. Individuals with this condition have intellectual disability and severe speech problems. They may also have an opening in the roof of the mouth, dental abnormalities, or other abnormalities of the head and face (craniofacial anomalies). Some of these mutations are deletions of large pieces of DNA that remove several genes, including *SATB2*. Other mutations add, remove, or rearrange smaller pieces of DNA within the *SATB2* gene. Still other mutations change single DNA building blocks (nucleotides) in the *SATB2* gene. It is likely that these genetic changes reduce the amount of functional SATB2 protein. Reduction of SATB2 function is thought to impair normal development of the brain and craniofacial structures, leading to intellectual disability, delayed speech, craniofacial anomalies, and other features of *SATB2*-associated syndrome.

**DHODH & HP:0000625**

HPO: Eyelid coloboma: The lateral segment of the lower eyelid is most commonly involved. As the milder forms of this finding are clearly subjective and no boundary of subjective and objective is defined, the term is considered subjective. The term eyelid coloboma has been replaced because the word coloboma should be used only for defects at the site of fusion of embryologic structures, which is not the case here. Modifiers to designate the location of the cleft may be added, such as lower and lateral.

Gene: At least 11 mutations in the *DHODH* gene have been found to cause Miller syndrome. Most of these mutations change single protein building blocks (amino acids) in dihydroorotate dehydrogenase, which likely impairs the enzyme's ability to function normally. It is unclear exactly how *DHODH* gene mutations lead to the signs and symptoms of Miller syndrome.

Children with Miller syndrome are born with underdeveloped cheek bones (malar hypoplasia) and a very small lower jaw ([micrognathia](https://ghr.nlm.nih.gov/art/large/micrognathia.jpeg)). They often have an opening in the roof of the mouth ([cleft palate](https://ghr.nlm.nih.gov/art/large/baby-with-cleft-palate.jpeg)) and/or a split in the upper lip ([cleft lip](https://ghr.nlm.nih.gov/art/large/baby-with-cleft-lip.jpeg)). These abnormalities frequently cause feeding problems in infants with Miller syndrome. The airway is usually restricted due to the micrognathia, which can lead to life-threatening breathing problems.

People with Miller syndrome often have eyes that slant downward, eyelids that turn out so the inner surface is exposed ([ectropion](https://ghr.nlm.nih.gov/art/large/ectropion.jpeg)), and a notch in the lower eyelids called an [eyelid coloboma](https://ghr.nlm.nih.gov/art/large/cleft-eyelid.jpeg).

**SATB2 & HP:0008872**

HPO: Feeding difficulties in infancy : *Impaired feeding performance of an infant as manifested by difficulties such as weak and ineffective sucking, brief bursts of sucking, and falling asleep during sucking. There may be difficulties with chewing or maintaining attention.*

Gene: The *SATB2* gene provides instructions for making a protein that helps control the development of certain body systems. The SATB2 protein attaches to special regions of DNA called matrix attachment regions (MARs). These regions help determine the structure of chromatin, which is the complex of DNA and proteins that packages DNA into chromosomes.

Mutations in the *SATB2* gene have been found to cause *SATB2*-associated syndrome. Individuals with this condition have intellectual disability and severe speech problems. They may also have an opening in the roof of the mouth, dental abnormalities, or other abnormalities of the head and face (craniofacial anomalies). Some of these mutations are deletions of large pieces of DNA that remove several genes, including *SATB2*

**LRP1 & HP:0000121**

HPO: Nephrocalcinosis: *Nephrocalcinosis is the deposition of calcium salts in renal parenchyma.* Nephrocalcinosis can be intratubular or interstitial, and can be diagnosed by means of a radiologic exam (plain radiographs, ultrasonograms, or computed tomography scans) or via microscopic examination of the renal tissues. The term nephrocalcinosis most often applies to a generalized increase in renal calcium content.

Gene: This gene encodes a member of the low-density lipoprotein receptor family of proteins. The encoded preproprotein is proteolytically processed by furin to generate 515 kDa and 85 kDa subunits that form the mature receptor (PMID: 8546712). This receptor is involved in several cellular processes, including intracellular signaling, lipid homeostasis, and clearance of apoptotic cells. In addition, the encoded protein is necessary for the alpha 2-macroglobulin-mediated clearance of secreted amyloid precursor protein and beta-amyloid, the main component of amyloid plaques found in Alzheimer patients.

Keratosis pilaris atrophicans (KPA): A group of rare genodermatoses characterized by keratotic follicular papules, variable degrees of inflammation, and secondary atrophic scarring. Most cases are associated with an atopic diathesis and keratosis pilaris on the extensor extremities. KPA is comprised of three distinct clinical subtypes: keratosis pilaris atrophicans faciei, atrophoderma vermiculatum, and keratosis follicularis spinulosa decalvans. Affected individuals may present with features overlapping the 3 subtypes. [MIM:604093]

**LRP1 & HP:0006530**

HPO: Interstitial pulmonary abnormality: *Abnormality of the lung parenchyma extending to the pulmonary interstitium and leading to diffuse pulmonary fibrosis.*

Gene: This gene encodes a member of the low-density lipoprotein receptor family of proteins. The encoded preproprotein is proteolytically processed by furin to generate 515 kDa and 85 kDa subunits that form the mature receptor (PMID: 8546712). This receptor is involved in several cellular processes, including intracellular signaling, lipid homeostasis, and clearance of apoptotic cells. In addition, the encoded protein is necessary for the alpha 2-macroglobulin-mediated clearance of secreted amyloid precursor protein and beta-amyloid, the main component of amyloid plaques found in Alzheimer patients.

Keratosis pilaris atrophicans (KPA): A group of rare genodermatoses characterized by keratotic follicular papules, variable degrees of inflammation, and secondary atrophic scarring. Most cases are associated with an atopic diathesis and keratosis pilaris on the extensor extremities. KPA is comprised of three distinct clinical subtypes: keratosis pilaris atrophicans faciei, atrophoderma vermiculatum, and keratosis follicularis spinulosa decalvans. Affected individuals may present with features overlapping the 3 subtypes. [MIM:604093]

**TMEM231 & HP:0000817**

HPO: Poor eye contact: Eye contact, define as the meeting of the gaze between two people during a conversation, is an important form of nonverbal communication.

Gene: This gene encodes a transmembrane protein, which is a component of the B9 complex involved in the formation of the diffusion barrier between the cilia and plasma membrane. Mutations in this gene cause Joubert syndrome (JBTS). Multiple alternatively spliced transcript variants have been found for this gene. [provided by RefSeq, Jan 2013]

Known Disease: Joubert syndrome 20 (JBTS20): A disorder presenting with cerebellar ataxia, oculomotor apraxia, hypotonia, neonatal breathing abnormalities and psychomotor delay. Neuroradiologically, it is characterized by cerebellar vermian hypoplasia/aplasia, thickened and reoriented superior cerebellar peduncles, and an abnormally large interpeduncular fossa, giving the appearance of a molar tooth on transaxial slices (molar tooth sign). Additional variable features include retinal dystrophy and renal disease. [MIM:614970]

Meckel syndrome 11 (MKS11): A disorder characterized by a combination of renal cysts and variably associated features including developmental anomalies of the central nervous system (typically encephalocele), hepatic ductal dysplasia and cysts, and polydactyly. [MIM:615397]

**GLB1 & HP:0008454**

HPO: Lumbar kyphosis: *Over curvature of the lumbar region*: *Rounded lower back, Lumbar gibbus deformity*

Gene: The *GLB1* gene provides instructions for producing an enzyme called beta-galactosidase (β-galactosidase). The *GLB1* gene also provides instructions for making the elastin-binding protein. On the cell surface, elastin-binding protein interacts with proteins called cathepsin A and neuraminidase 1 to form the elastin receptor complex. This receptor complex plays a role in the formation of elastic fibers, which are a component of the connective tissue that forms the body's supportive framework.

Disease: More than 80 mutations in the *GLB1* gene have been found to cause GM1 gangliosidosis. Most mutations change single DNA building blocks (nucleotides) in the *GLB1* gene. These mutations often affect the production of both β-galactosidase and elastin-binding protein.

More than 10 mutations in the *GLB1* gene have been found to cause mucopolysaccharidosis type IV (MPS IV). The lack of β-galactosidase activity leads to the accumulation of keratan sulfate within lysosomes. Because keratan sulfate is predominantly found in cartilage and the cornea, the buildup of this substance causes skeletal abnormalities and cloudy corneas

**IFT140 & HP:0001636**

HPO: Tetralogy of Fallot: *A congenital cardiac malformation comprising pulmonary stenosis, overriding aorta, ventricular septum defect, and right ventricular hypertrophy. The diagnosis of TOF is made if at least three of the four above mentioned features are present.*

Gene: The *IFT140* gene provides instructions for making a protein that is involved in the formation and maintenance of cilia

Disease: Mainzer-Saldino syndrome, a disorder characterized by kidney disease, eye problems, and skeletal abnormalities. Mutations in the *IFT140* gene that cause Mainzer-Saldino syndrome may change the shape of the IFT140 protein or its interactions with other IFT proteins, likely impairing the assembly of IFT-A and the development or maintenance of cilia.

Asphyxiating thoracic dystrophy, also known as Jeune syndrome, is an inherited disorder of bone growth characterized by a narrow chest, short ribs, shortened bones in the arms and legs, short stature, and extra [fingers](https://ghr.nlm.nih.gov/art/large/postaxial-polydactyly-hand.jpeg) and [toes](https://ghr.nlm.nih.gov/art/large/preaxial-polydactyly-foot.jpeg) (polydactyly). Additional skeletal abnormalities can include unusually shaped collarbones (clavicles) and pelvic bones, and and cone-shaped [ends of the long bones](https://ghr.nlm.nih.gov/art/large/parts-of-a-normal-bone.jpeg) in the arms and legs. Many infants with this condition are born with an extremely narrow, bell-shaped chest that can restrict the growth and expansion of the lungs. Life-threatening problems with breathing result, and people with asphyxiating thoracic dystrophy may live only into infancy or early childhood. However, in people who survive beyond the first few years, the narrow chest and related breathing problems can improve with age.

**DHODH & HP:0000202**

HPO: Oral cleft: *The presence of a cleft in the oral cavity, the two main types of which are cleft lip and cleft palate. In cleft lip, there is the congenital failure of the maxillary and median nasal processes to fuse, forming a groove or fissure in the lip. In cleft palate, there is a congenital failure of the palate to fuse properly, forming a grooved depression or fissure in the roof of the mouth. Clefts of the lip and palate can occur individually or together. It is preferable to code each defect separately.*

Gene: At least 11 mutations in the *DHODH* gene have been found to cause Miller syndrome. Most of these mutations change single protein building blocks (amino acids) in dihydroorotate dehydrogenase, which likely impairs the enzyme's ability to function normally. It is unclear exactly how *DHODH* gene mutations lead to the signs and symptoms of Miller syndrome.

Children with Miller syndrome are born with underdeveloped cheek bones (malar hypoplasia) and a very small lower jaw ([micrognathia](https://ghr.nlm.nih.gov/art/large/micrognathia.jpeg)). They often have an opening in the roof of the mouth ([cleft palate](https://ghr.nlm.nih.gov/art/large/baby-with-cleft-palate.jpeg)) and/or a split in the upper lip ([cleft lip](https://ghr.nlm.nih.gov/art/large/baby-with-cleft-lip.jpeg)). These abnormalities frequently cause feeding problems in infants with Miller syndrome. The airway is usually restricted due to the micrognathia, which can lead to life-threatening breathing problems.

People with Miller syndrome often have eyes that slant downward, eyelids that turn out so the inner surface is exposed ([ectropion](https://ghr.nlm.nih.gov/art/large/ectropion.jpeg)), and a notch in the lower eyelids called an [eyelid coloboma](https://ghr.nlm.nih.gov/art/large/cleft-eyelid.jpeg).

**DHODH & HP:0000448**

HPO: Prominent nose: *Distance between subnasale and pronasale more than two standard deviations above the mean, or alternatively, an apparently increased anterior protrusion of the nasal tip*

Gene: At least 11 mutations in the *DHODH* gene have been found to cause Miller syndrome. Most of these mutations change single protein building blocks (amino acids) in dihydroorotate dehydrogenase, which likely impairs the enzyme's ability to function normally. It is unclear exactly how *DHODH* gene mutations lead to the signs and symptoms of Miller syndrome.

Children with Miller syndrome are born with underdeveloped cheek bones (malar hypoplasia) and a very small lower jaw ([micrognathia](https://ghr.nlm.nih.gov/art/large/micrognathia.jpeg)). They often have an opening in the roof of the mouth ([cleft palate](https://ghr.nlm.nih.gov/art/large/baby-with-cleft-palate.jpeg)) and/or a split in the upper lip ([cleft lip](https://ghr.nlm.nih.gov/art/large/baby-with-cleft-lip.jpeg)). These abnormalities frequently cause feeding problems in infants with Miller syndrome. The airway is usually restricted due to the micrognathia, which can lead to life-threatening breathing problems.

People with Miller syndrome often have eyes that slant downward, eyelids that turn out so the inner surface is exposed ([ectropion](https://ghr.nlm.nih.gov/art/large/ectropion.jpeg)), and a notch in the lower eyelids called an [eyelid coloboma](https://ghr.nlm.nih.gov/art/large/cleft-eyelid.jpeg).

**SF3B4 & HP:0011800**

HPO: Midface retrusion: *Posterior positions and/or vertical shortening of the infraorbital and perialar regions, or increased concavity of the face and/or reduced nasolabial angle. Posterior positions and/or vertical shortening of the infraorbital and perialar regions, or increased concavity of the face and/or reduced nasolabial angle.*

Gene: The *SF3B4* gene provides instructions for making the SAP49 protein, which is part of a complex called a spliceosome.

The SAP49 protein may also be involved in a chemical signaling pathway known as the bone morphogenic protein (BMP) pathway. This signaling pathway regulates various cellular processes and is involved in the growth of cells. The SAP49 protein is particularly important for the maturation of cells that build bones and cartilage (osteoblasts and chondrocytes).

More than 30 mutations in the *SF3B4* gene have been found to cause Nager syndrome, which is primarily characterized by abnormalities of the face, hands, and arms, such as underdeveloped cheek bones (malar hypoplasia)

**SS18L1 & HP:0002190**

HPO: Choroid plexus cyst : Choroid plexus cysts can be observed on prenatal ultrasound examinations and are associated with a weakly increased risk for fetal chromosome abnormalities such as trisomy 18.

Gene: The SS18L1 gene encodes a calcium-responsive transactivator (CREST) that is an essential subunit of a neuron-specific chromatin-remodeling complex (nBAF)

Unknown significance (OMIM 3 cases).

**SS18L1 & HP:0002190**

HPO: Abnormal intestine morphology

Gene: The SS18L1 gene encodes a calcium-responsive transactivator (CREST) that is an essential subunit of a neuron-specific chromatin-remodeling complex (nBAF)

Unknown significance (OMIM 3 cases).

**COG5 & HP:0100512**

HPO: Low levels of vitamin D

Gene: The *COG5* gene provides instructions for making a protein called component of oligomeric Golgi complex 5 (COG5). As its name suggests, COG5 is one piece of a group of proteins known as the conserved oligomeric Golgi (COG) complex.

Eight mutations in the *COG5* gene are known to cause *COG5*-congenital disorder of glycosylation (*COG5*-CDG). This condition often leads to developmental delay and intellectual disability and causes other abnormalities.

**COG5 & HP:0010522**

HPO: Dyslexia

Gene: The *COG5* gene provides instructions for making a protein called component of oligomeric Golgi complex 5 (COG5). As its name suggests, COG5 is one piece of a group of proteins known as the conserved oligomeric Golgi (COG) complex.

Eight mutations in the *COG5* gene are known to cause *COG5*-congenital disorder of glycosylation (*COG5*-CDG). This condition often leads to developmental delay and intellectual disability and causes other abnormalities.