

## Filtering for Compound Heterozygous Sequence Variants in Non-Consanguineous Pedigrees

The identification of disease-causing mutations in next-generation sequencing (NGS) data requires efficient filtering techniques. For patients with rare recessive diseases, compound heterozygosity of pathogenic mutations is the likeliest inheritance model if the parents are non-consanguineous.

GeneTalk offers an efficient filter for compound heterozygous sequence variants. Open the *Filter* menu and select a VCF file that is listed under *Collections* (i.e., the file contains at least a set of three individuals and is supplied with pedigree information)\*. Now, you will find the *compound heterozygous* filter option in the *Inheritance Filter* tab.

### Background

In recessive genetic disorders all copies of a certain gene are malfunctioning. For autosomal recessive genes this means that the maternally as well as the paternally inherited copy of a gene harbors a pathogenic mutation. The occurrence of a mutation is a random process and many different pathogenic mutations have arisen from recessive genes in the human population over time. It means the

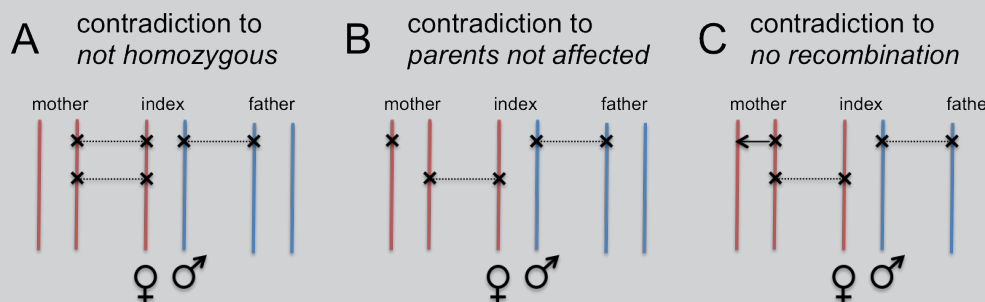
lower the kinship of the parents the higher is the chance that two different mutant alleles of the disease-causing gene are present. Hence, this can be translated into a simple rule of thumb: If the parents are non-consanguineous, compound heterozygosity is the likeliest explanation for a recessive disease.

A challenge in filtering sequence variants for compound heterozygotes is that one has to figure out whether the two heterozygous variants affect different copies of a gene or the same copy. Usually, that cannot be determined from a single DNA sequence, when the read length is smaller than the distance between the variants. However, when exome data of more than one family member is available, one can exclude certain variants based on rules of Mendelian inheritance (see box).


















## Recessive Compound Heterozygous Filter Rules

To be transmitted in a compound heterozygous mode from parents to the patient, a variant must meet the following conditions:

1. The variant is heterozygous in all affected individuals.
2. The variant is not homozygous in any of the unaffected individuals.
3. The variant is heterozygous in exactly one of the parents if it is heterozygous in an affected child.
  - a. The variant is not heterozygous in both parents.
  - b. The variant is present in at least one parent.
4. There is at least a second heterozygous variant in the same gene in each of the affected individuals.
5. There is at least one variant transmitted from the paternal side and one transmitted from the maternal side.



\* Requires a GeneTalk premium license.

Chrom	Pos	dbSNP ID	Gene	Ref	465 (Affected)	466 (Unaffected)	467 (Unaffected)	Effect	Medical Relevance	Scientific Evidence	
1	145296451	 rs140256918	NBPF10 	G	C/G	C/G	G/G	missense			
1	145368518	 rs61813437	NBPF10 	C	C/T	C/C	C/T	missense			
9	35090263	 rs142164373	PIGO 	G	A/G	A/G	G/G	missense			
9	35091529		PIGO 	G	G/GG	G/G	G/GG	frameshift-insertion			






## Example

An analysis of exome data from a family with two siblings affected by Mabry syndrome (also known as hyperphosphatasia with mental retardation syndrome) demonstrates how effective the inheritance filter reduced the amount of sequence variants. Starting with 27,584 variants, the data was filtered for genotype frequency (0.1%) and effects on protein level yielding only 1446 variants. The compound heterozygous filter reduces this set to just six variants in three distinct genes, *NBPF10*, *MUC4*, and *PIGO*. A quick assessment of the medical relevance and scientific evidence of the variant calls reveals *PIGO* as the disease candidate gene and the mutations were confirmed as the causative mutations of the disorder.

## Conclusion

The rule set, on which the filter is based, is comprehensive for analyzing multiple samples and advances the prioritization of compound heterozygous variants. It shows that filtering for compound heterozygous mutations is an effective means in identifying disease candidate genes especially when several family members are available for the analysis. In a trio analysis, for example, with exome data of the parents and one affected child, just about a dozen of variants in candidate genes remain after applying the filter. This manageable number of remaining genes can then be assessed based on the expertise of the investigator or further prioritized using the tools in GeneTalk.

### Annotation

<p><b>Position</b> chr9 : 35090263</p> <p><b>Genotype</b> A/G</p> <p><b>Gene</b> <a href="#">PIGO</a> (Entrez-ID: <a href="#">84720</a>)</p> <p><b>HGVS DNA code</b> NM_032634.3:c.2869C&gt;T</p> <p><b>HGVS protein code</b> NP_116023.2:p.F957L</p> <p><b>Author:</b> <a href="#">petkraw</a> </p> <p><b>Text:</b> This variant impairs the function of PIGO. In a homozygous or compound heterozygous state this variant will cause hyperphosphatasia with mental retardation syndrome also known as Mabry syndrome</p>	<p><b>Link</b>  <a href="http://ghr.nlm.nih.gov/condition/mabr...">http://ghr.nlm.nih.gov/condition/mabr...</a></p> <p><b>OMIM-ID</b>  <a href="#">614749</a></p> <p><b>Pubmed Links</b>  <a href="#">Pubmed Article 22683086</a></p> <p><b>MutationTaster</b> </p> <p style="text-align: right; font-size: small;">Created at June 06, 2012 12:25, Last update January 31, 2014 09:19</p>
---	--

## References:

Krawitz P.M et al., Mutations in PIGO, a member of the GPI-anchor-synthesis pathway, cause hyperphosphatasia with mental retardation. *Am J Hum Genet*, 2012, Kamphans T. et al., Filtering for Compound Heterozygous Sequence Variants in Non-Consanguineous Pedigrees. *PLoS ONE*, 2013