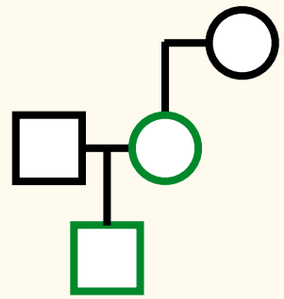
Classifying inherited variants

Elizabeth K. Ruzzo
Laura Perez-Cano

Pre-analysis



Seq data (n=4)
*Phase-able (n=2)

1 Extract list of affected and unaffected phase-able* individuals

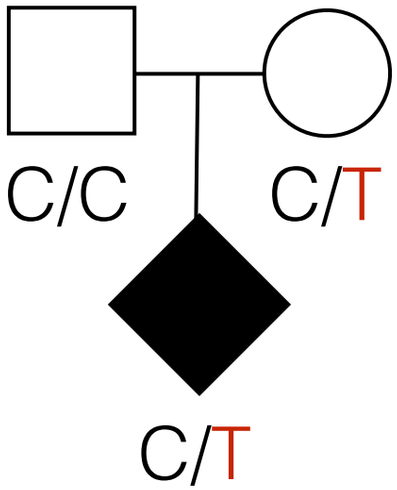
2 Define genotypes: 0 (Hom ref), 1 (Het), 2 (Hom alt), 3 (QC failure) and missing (./.)

3 Determine inheritance of variants (8 possible categories)

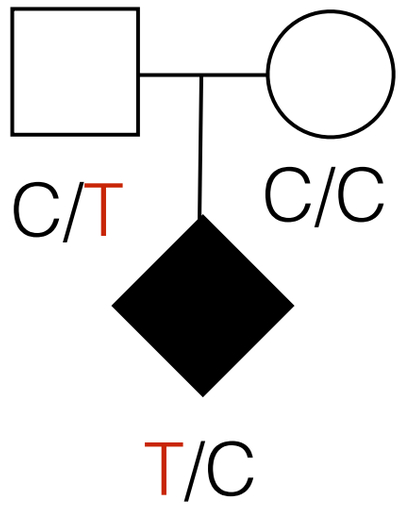
4 Counts for each variant and inheritance type

Chr	**Pos-ition**	**Ref**	**Alt**	**Info**	De Novo	num_aff_dn	ids_aff_dn	num_unaff_dn	ids_unaff_dn
					Newly homozygous	num_aff_homo	ids_aff_homo	num_unaff_homo	ids_unaff_homo
					Hemizygous	…	…	…	…
					From mother	…	…	…	…
					From father	…	…	…	…
					Unknown phase	…	…	…	…
					Missing subjects	…	…	…	…
					Uncertain	…	…	…	…

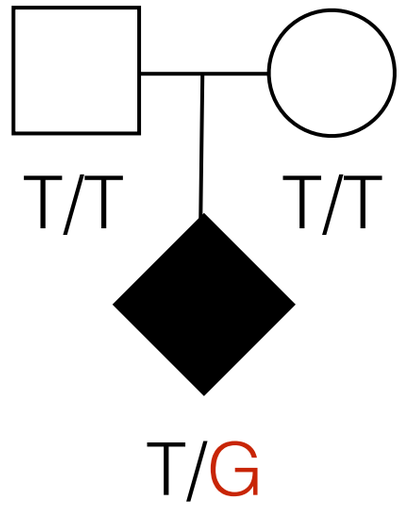
Maternally inherited



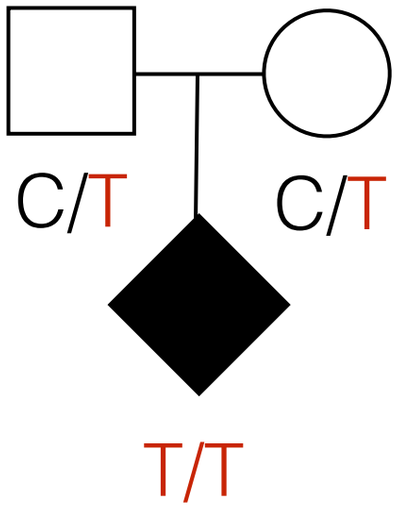
Paternally inherited



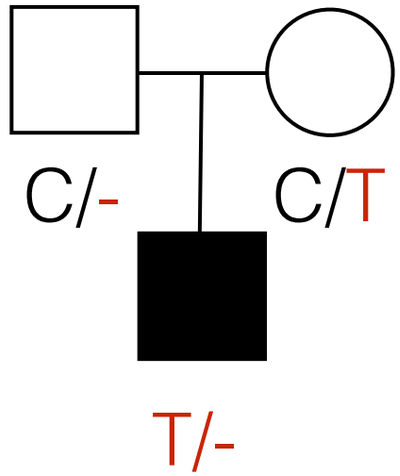
De Novo



Newly Homozygous



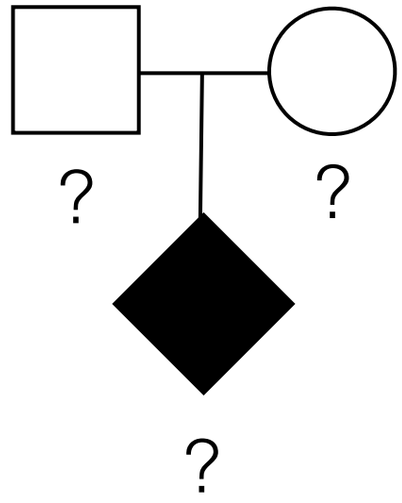
Newly Hemizygous (ChrX)



6,7,8

Other

Missing
Uncertain
Unknown phase



Critical Assumption

- If a variant is present and passed specified QC thresholds, then the assigned genotype is the correct genotype
- Recommended QC thresholds: -d 10 -a 0 -r 0.2 -g 25
  - Depth (-d): ≥10x
  - Alt allele depth (-a): ≥0 (min number of reads required)
  - Ratio of alt allele reads/total reads (-r): ≥0.2
  - Genotype quality (-g): ≥25

Output notes

- If a variant is missing or homozygous reference in all phase-able individuals, this variant is not output. However, it may be present somewhere, for example, only in the unaffected parents and not transmitted.

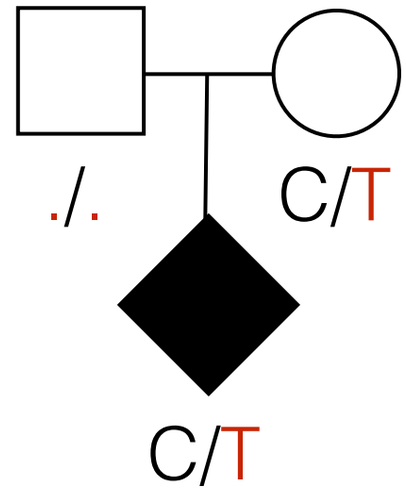
- If a variant is output, the only samples not listed in one of the eight categories are samples which are homozygous reference (and pass QC) at the variant site.

Rules

- De Novo
  - Example: subject is heterozygous (0/1) or homozygous (1/1) and both parents are hom ref (0/0)
  - Thinking is if homozygous perhaps low coverage (only Mendelian error accepted)

Rules

- Unknown-phase: Unclear where variant was inherited from.
    - Example: subject is heterozygous (0/1) and one parent also het (0/1) and one parent is also het or hom (0/1 or 1/1)
    - Example 2: subject is heterozygous (0/1) and one parent is het (0/1) and one parent is missing (./.)



- Uncertain:
    1. Variant could be in more than one category
    2. Incomplete family
        - Example: If mother is missing, a variant in a son on the X-chromosome could either be “newly hemizygous” or “maternally inherited”
    3. QC failure
        - Example: Mother was sequenced but the variant site has low coverage.
    4. Mendelian Error
        - Includes het variants on male sex chromosomes
        - Exception: homozygous de novo variants
    5. Homozygous (and also homozygous in parent(s))