Genotype imputation

The scientific art of guessing stuff

Lecturer profile

Giannos Louloudis, Research Assistant at CPR

Novo Nordisk Foundation Center for Protein Research



Copenhagen, Denmark

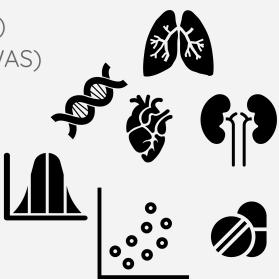
Brunak group



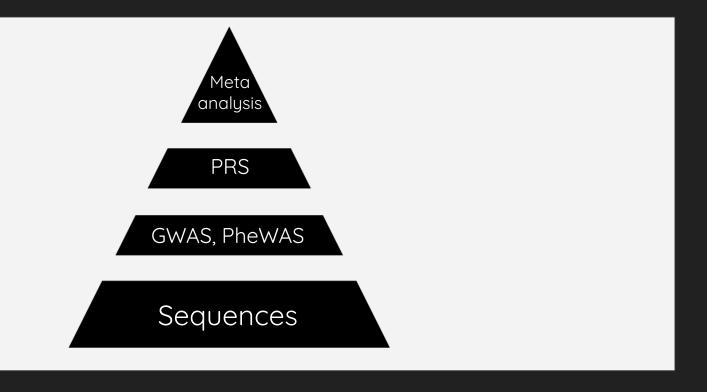
Research interests: Translational Disease Systems Biology

- Genome-wide association studies (GWAS)
- Phenome-wide association studies (PheWAS)
- Polygenic Risk Scores (PRS)

- Survival analysis
- Drug repositioning



Basic structure



Whole Genome Sequencing vs Global Sequencing Array

Theory vs Real world

Uses and applications

Whole Genome Sequencing (WGS)

The process of determining the entirety of an individual's genome sequence.

Chr 1 Chr 2 Chr 3 Chr 4 Chr 5 Chr 6 Chr 7 Chr 8 Chr 9

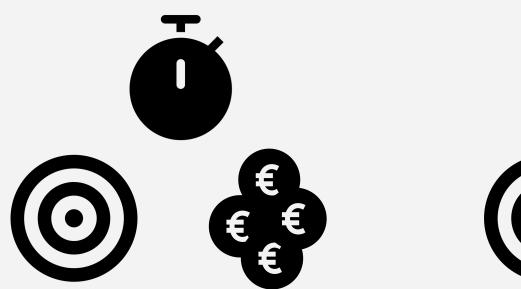
Global Sequencing Array (GSA)

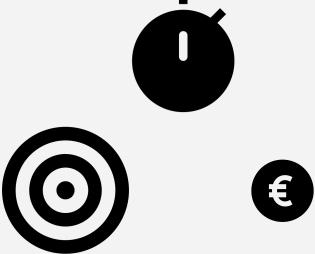
Array genotyping a (known) part of an individual's genome. *Specific SNPs*

```
Chr 1
Chr 2
Chr 3
Chr 4
Chr 5
Chr 6
Chr 7
Chr 8
Chr 9
```

Whole Genome Sequencing (WGS)

Global Sequencing Array (GSA)





Uses and applications

Global Sequencing Array (GSA) Whole Genome Sequencing (WGS)





Bioinformatics file naming.

>4466584.3|G1E3M3B04IX1IW|Greengenes|263471 16S ribosomal RNA [Microbacterium oxydans] qqctATAATTTGTAAATTTCTTGAGATAGAATCATTCGTATTGAATGAGGTCAAATTCTC TAAACTGATTAAGAAGTATAATACTTAGATGCGAGTTATTGCATCACTTAACGGAGAGTT TGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAACGTGAAG TCTGAATTGAGTACTTCGGTATGATATTTGGGTGGAAAGTGGCGGACGGGTGAGTAACAC GTGGGTAACCTGCCTCGAAGTGGGGACAACCATTGGAAACGATGGCTAATACCGCATAGT TCTTTAGATGCATGAGCATTTATAGATAAAACTCTGGTGCTTCGAGAGGGGTCTGCGTCC GATTAGTTAGTTGGTGGGTAAAGGCCTACCAAGACGATGATCGGTAGCTGGTCTGAGAGG ACGATCAGTCACACGGGAACTGAGACACGGTCCagtcgtgggagacaaggcacacagggg atagannnnn >4466584.3|G1E3M3B04IX1IW|Greengenes|265788 16S ribosomal RNA [Microbacterium oxydans] gactATAATTTGTAAATTTCTTGAGATAGAATCATTCGTATTGAATGAGGTCAAATTCTC TAAACTGATTAAGAAGTATAATACTTAGATGCGAGTTATTGCATCACTTAACGGAGAGTT TGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAACGTGAAG TCTGAATTGAGTACTTCGGTATGATATTTGGGTGGAAAGTGGCGGACGGGTGAGTAACAC GTGGGTAACCTGCCTCGAAGTGGGGACAACCATTGGAAACGATGGCTAATACCGCATAGT TCTTTAGATGCATGAGCATTTATAGATAAAACTCTGGTGCTTCGAGAGGGGTCTGCGTCC GATTAGTTAGTTGGTGGGTAAAGGCCTACCAAGACGATGATCGGTAGCTGGTCTGAGAGG ACGATCAGTCACACGGGAACTGAGACACGGTCCagtcgtgggagacaaggcacacagggg ataaannnnn

Fasta files

>read no 1 >read no 2 GCAGCGTGAGCGCCATCATGGGCAACCCCCAGGTGAAGGCCCACGGCAAGA >read no 3 GGGAGACACCCGCACGTGTGGCCCGCATGTATGCTGAGCTCTTCCGCGGAT >read no 4 TTTGCCCCGCATCGAGCGGGCTGTGCGGGAAATCCTTCTGGCTGTAGGCGA >read no 5 CCTGTGGGGCAAGGTGAACCCCGTGGAGATCGGCGCGAGAGCCTGGCCAG >read no 6 GAGGAGGCCAGGATCCACCAGAGGAAGGGCCTGCTGTGGTTCATCCCCGC >read_no_7 CTGCACAGCGACTACAACCTGACCTGGTACAGGAACGGCAGCAACATGCCC >read no 8 GTGCTGGGCCTGGCCATCAGCCACTTCCTGCTGGAGCAGTTCCCCGACTAC >read no 9 AACCTGGGCGAGTACCTGCTGCTGGGCAAGGGCGAGGAGATGACCGGCGGC >read no 10 GTTCCCCGACTACAACGAGGGCGAGCTGAGCAGGCTGAGGAGCGCCATCGT >read no 11 CTTCAGCAAGTTCGGCGACCTGAGCAGCGTGAGCGCCATCATGGGCAACCC >read no 12 ACCAGAGGAAGGGCCTGCTGTGGTTCATCCCCGCCGCCCTGGAGGACAGCG >read no 13 AAGGGCGAGGAGATGACCGGCGGCAGGAGGAAGGCCAGCCTGCTGGCCGAC

Variant Call Format

##fileformat=VCFv4.3

```
##fileDate=20230307
##source=PLINKv2.00
##contig=<ID=1>
##contia=<ID=2>
##contig=<ID=3>
##contig=<ID=4>
##contia=<ID=5>
##contig=<ID=6>
##contig=<ID=7>
##contig=<ID=8>
##contig=<ID=9>
##contig=<ID=10>
##contig=<ID=11>
##contig=<ID=12>
##contig=<ID=13>
##contig=<ID=14>
##contig=<ID=15>
##contig=<ID=16>
##contig=<ID=17>
##contig=<ID=18>
##contig=<ID=19>
##contig=<ID=20>
##contig=<ID=21>
##contig=<ID=22>
##contig=<ID=Y>
##INFO=<ID=PR, Number=0, Type=Flag, Description="Provisional reference allele, may not be based on real reference genome">
##FORMAT=<ID=GT, Number=1, Type=String, Description="Genotype">
                                                                                                                      HGDP01361 HGDP01361
#CHROM POS ID
                               ALT
                                       QUAL
                                               FILTER INFO
                                                               FORMAT
                                                                      HGDP00521 HGDP00521
                                                                                               HGDP00533 HGDP00533
                                                                                                                                              HGDP01372 HGDP01372
                                                                                                                                                                      HGDP00667 HGDP00667
                                                                                                                                                                                             HGDP01066 HGDP01066
                                                                                                                                                                                                                     HGDP01078 HGDP01078
                                                                                                                                                                                                                                            HGDP01>
        768448 rs12562034
                                                                               9/9
                                                                                      0/0
                                                                                              9/9
                                                                                                       0/1
                                                                                                              0/0
                                                                                                                      0/0
                                                                                                                              9/9
                                                                                                                                     0/0
                                                                                                                                              9/9
                                                                                                                                                      0/0
                                                                                                                                                             9/1
                                                                                                                                                                     0/0
                                                                                                                                                                             9/1
                                                                                                                                                                                     9/9
                                                                                                                                                                                             9/9
                                                                                                                                                                                                     0/0
                                                                                                                                                                                                             9/1
                                                                                                                                                                                                                     9/9
                                                                                                                                                                                                                             0/1
                                                                                                                                                                                                                                    0/0
                                                                                                                                                                                                                                            0/0
        1005806 rs3934834
                                                                               0/0
                                                                                      0/1
                                                                                                              0/0
                                                                                                                      0/1
                                                                                                                              0/0
                                                                                                                                     0/1
                                                                                                                                              0/1
                                                                                                                                                      0/0
                                                                                                                                                             0/1
                                                                                                                                                                     0/0
                                                                                                                                                                             0/0
                                                                                                                                                                                     0/0
                                                                                                                                                                                                             0/0
                                                                                                                                                                                                                     0/0
                                                                                                                                                                                                                             0/1
                                                                                                                                                                                                                                    0/0
                                                                                                                                                                                                                                            0/0
                                                               PR
                                                                                                      0/0
                                                                                                                                                                                                     0/1
       1018704 rs9442372
                                                               PR
                                                                               0/0
                                                                                      0/1
                                                                                                       1/1
                                                                                                              0/0
                                                                                                                      0/1
                                                                                                                              0/1
                                                                                                                                      0/1
                                                                                                                                                      0/0
                                                                                                                                                                     0/0
                                                                                                                                                                             0/0
                                                                                                                                                                                     0/0
                                                                                                                                                                                                     0/1
                                                                                                                                                                                                             1/1
                                                                                                                                                                                                                     0/0
                                                                                                                                                                                                                             0/1
                                                                                                                                                                                                                                    0/1
                                                                                                                                                                                                                                            0/0
        1021695 rs9442398
                                                               PR
                                                                               0/0
                                                                                      0/0
                                                                                                       1/1
                                                                                                              0/0
                                                                                                                      0/0
                                                                                                                              0/0
                                                                                                                                      0/0
                                                                                                                                              0/1
                                                                                                                                                      0/0
                                                                                                                                                             0/1
                                                                                                                                                                     0/0
                                                                                                                                                                              0/0
                                                                                                                                                                                     0/0
                                                                                                                                                                                                                     0/0
                                                                                                                                                                                                                             0/0
                                                                                                                                                                                                                                    0/1
                                                                                                                                                                                                                                            0/0
                                                                                                                                                                                                     0/0
        1030565 rs6687776
                                                               PR
                                                                       GT
                                                                               0/0
                                                                                      0/0
                                                                                              0/0
                                                                                                       0/1
                                                                                                              0/0
                                                                                                                      0/0
                                                                                                                              0/0
                                                                                                                                      0/0
                                                                                                                                              0/1
                                                                                                                                                      0/0
                                                                                                                                                             0/1
                                                                                                                                                                     0/1
                                                                                                                                                                             0/1
                                                                                                                                                                                     0/0
                                                                                                                                                                                             0/0
                                                                                                                                                                                                     0/1
                                                                                                                                                                                                             0/0
                                                                                                                                                                                                                             0/1
                                                                                                                                                                                                                                    0/0
                                                                                                                                                                                                                                            0/0
        1031540 rs9651273
                                                               PR
                                                                       GT
                                                                               0/0
                                                                                      0/1
                                                                                                       0/0
                                                                                                              0/0
                                                                                                                      0/1
                                                                                                                              0/1
                                                                                                                                      0/1
                                                                                                                                              0/1
                                                                                                                                                      0/0
                                                                                                                                                             0/1
                                                                                                                                                                     0/0
                                                                                                                                                                              0/0
                                                                                                                                                                                     0/0
                                                                                                                                                                                                     0/0
                                                                                                                                                                                                                     0/0
                                                                                                                                                                                                                             0/0
                                                                                                                                                                                                                                    0/1
                                                                                                                                                                                                                                            0/0
        1048955 rs4970405
                                                               PR
                                                                               0/0
                                                                                      0/0
                                                                                               0/0
                                                                                                       0/1
                                                                                                              0/0
                                                                                                                      0/0
                                                                                                                              0/0
                                                                                                                                      0/0
                                                                                                                                              0/1
                                                                                                                                                      0/0
                                                                                                                                                             0/1
                                                                                                                                                                     0/1
                                                                                                                                                                              0/0
                                                                                                                                                                                     0/0
                                                                                                                                                                                             0/0
                                                                                                                                                                                                     0/1
                                                                                                                                                                                                             0/0
                                                                                                                                                                                                                     0/1
                                                                                                                                                                                                                             0/1
                                                                                                                                                                                                                                    0/0
                                                                                                                                                                                                                                            0/0
        1049950 rs12726255
                                                               PR
                                                                       GT
                                                                               0/0
                                                                                      0/0
                                                                                               0/0
                                                                                                       0/1
                                                                                                              0/0
                                                                                                                      0/0
                                                                                                                              0/0
                                                                                                                                      0/0
                                                                                                                                              0/1
                                                                                                                                                      0/0
                                                                                                                                                             0/1
                                                                                                                                                                     0/1
                                                                                                                                                                              0/0
                                                                                                                                                                                     0/0
                                                                                                                                                                                             0/0
                                                                                                                                                                                                     0/1
                                                                                                                                                                                                             0/0
                                                                                                                                                                                                                     0/1
                                                                                                                                                                                                                             0/1
                                                                                                                                                                                                                                    0/0
                                                                                                                                                                                                                                            0/0
        1060235 rs7540009
                                                               PR
                                                                       GT
                                                                               0/0
                                                                                      9/9
                                                                                               0/0
                                                                                                       0/0
                                                                                                              9/9
                                                                                                                      0/0
                                                                                                                              0/0
                                                                                                                                      0/0
                                                                                                                                              0/0
                                                                                                                                                      0/0
                                                                                                                                                             0/0
                                                                                                                                                                      0/0
                                                                                                                                                                              0/0
                                                                                                                                                                                     0/0
                                                                                                                                                                                             0/0
                                                                                                                                                                                                     0/0
                                                                                                                                                                                                             0/0
                                                                                                                                                                                                                     0/0
                                                                                                                                                                                                                             0/0
                                                                                                                                                                                                                                    9/9
                                                                                                                                                                                                                                            0/0
        1061166 rs11807848
                                                               PR
                                                                               0/0
                                                                                      0/1
                                                                                                              0/0
                                                                                                                      0/1
                                                                                                                              0/1
                                                                                                                                      0/1
                                                                                                                                              0/1
                                                                                                                                                      0/1
                                                                                                                                                                     0/1
                                                                                                                                                                              0/0
                                                                                                                                                                                     0/0
                                                                                                                                                                                                             0/1
                                                                                                                                                                                                                     0/1
                                                                                                                                                                                                                             0/1
                                                                                                                                                                                                                                    0/0
                                                                                                                                                                                                                                            0/0
                                                                                                       0/1
                                                                                                                                                                                                     0/1
                                                               PR
                                                                               0/0
                                                                                      0/1
                                                                                                                              0/1
                                                                                                                                                      0/1
                                                                                                                                                                     0/1
                                                                                                                                                                              0/0
                                                                                                                                                                                     0/0
                                                                                                                                                                                                             0/1
                                                                                                                                                                                                                     0/1
                                                                                                                                                                                                                                    0/0
                                                                                                                                                                                                                                            0/0
        1062638 rs9442373
                                                                                                      0/1
                                                                                                              0/1
                                                                                                                      0/1
                                                                                                                                      0/1
                                                                                                                                              0/1
                                                                                                                                                                                                                             0/1
        1064979 rs2298217
                                                               PR
                                                                               0/0
                                                                                      0/0
                                                                                              9/9
                                                                                                      0/1
                                                                                                              0/1
                                                                                                                      0/0
                                                                                                                              0/0
                                                                                                                                      0/0
                                                                                                                                              0/1
                                                                                                                                                      0/0
                                                                                                                                                             9/1
                                                                                                                                                                     0/1
                                                                                                                                                                              0/0
                                                                                                                                                                                     0/0
                                                                                                                                                                                             0/0
                                                                                                                                                                                                     0/1
                                                                                                                                                                                                             9/9
                                                                                                                                                                                                                     9/1
                                                                                                                                                                                                                             0/1
                                                                                                                                                                                                                                    0/0
                                                                                                                                                                                                                                            0/0
        1066029 rs12145826
                                                               PR
                                                                               0/1
                                                                                      0/1
                                                                                              0/0
                                                                                                                      0/0
                                                                                                                              0/1
                                                                                                                                      0/1
                                                                                                                                              0/0
                                                                                                                                                      0/1
                                                                                                                                                                     0/0
                                                                                                                                                                              0/0
                                                                                                                                                                                     0/0
                                                                                                                                                                                                             0/0
                                                                                                                                                                                                                     0/0
                                                                                                                                                                                                                                    0/0
                                                                                                                                                                                                                                            0/0
                                                                                                      0/0
                                                                                                              0/0
                                                                                                                                                             0/0
                                                                                                                                                                                                     0/1
                                                                                                                                                                                                                             0/0
                                                               PR
                                                                               0/0
                                                                                              0/0
                                                                                                                                                                     0/0
                                                                                                                                                                             0/0
                                                                                                                                                                                     0/0
                                                                                                                                                                                             0/0
                                                                                                                                                                                                                     0/0
                                                                                                                                                                                                                                    0/0
                                                                                                                                                                                                                                            0/0
        1077064 rs4970357
                                                                                      0/0
                                                                                                      9/9
                                                                                                                      0/0
                                                                                                                              0/0
                                                                                                                                      0/0
                                                                                                                                              0/0
                                                                                                                                                      0/0
                                                                                                                                                             0/1
                                                                                                                                                                                                     0/1
                                                                                                                                                                                                             0/0
                                                                                                                                                                                                                             0/1
                                                                                              0/0
                                                                                                                                                                                     0/0
                                                                                                                                                                                             0/0
                                                                                                                                                                                                                     0/0
                                                                                                                                                                                                                                    0/0
                                                                                                                                                                                                                                            0/0
        1087683 rs9442380
                                                                               0/0
                                                                                      0/0
                                                                                                       0/0
                                                                                                              1/1
                                                                                                                      0/0
                                                                                                                              0/0
                                                                                                                                      0/0
                                                                                                                                              0/0
                                                                                                                                                      0/0
                                                                                                                                                             0/1
                                                                                                                                                                     0/0
                                                                                                                                                                              0/0
                                                                                                                                                                                                     0/1
                                                                                                                                                                                                             0/0
                                                                                                                                                                                                                             0/1
                                                               PR
        1090557 rs7553429
                                                                       GT
                                                                               0/0
                                                                                      0/0
                                                                                              0/0
                                                                                                      0/0
                                                                                                              0/0
                                                                                                                      0/0
                                                                                                                              0/0
                                                                                                                                      0/0
                                                                                                                                              0/0
                                                                                                                                                      0/0
                                                                                                                                                             0/0
                                                                                                                                                                     0/0
                                                                                                                                                                             0/0
                                                                                                                                                                                     0/0
                                                                                                                                                                                             0/0
                                                                                                                                                                                                     0/0
                                                                                                                                                                                                             0/1
                                                                                                                                                                                                                     0/0
                                                                                                                                                                                                                             0/0
                                                                                                                                                                                                                                    0/0
                                                                                                                                                                                                                                            0/0
        1094738 rs4970362
                                                               PR
                                                                       GT
                                                                               0/0
                                                                                      0/0
                                                                                              0/1
                                                                                                       0/0
                                                                                                                      0/0
                                                                                                                              0/0
                                                                                                                                      0/0
                                                                                                                                              0/0
                                                                                                                                                      0/0
                                                                                                                                                                     0/0
                                                                                                                                                                              0/0
                                                                                                                                                                                     0/0
                                                                                                                                                                                             0/0
                                                                                                                                                                                                     0/1
                                                                                                                                                                                                             0/0
                                                                                                                                                                                                                     0/0
                                                                                                                                                                                                                             0/1
                                                                                                                                                                                                                                    0/0
                                                                                                                                                                                                                                            0/1
                                                               PR
        1099342 rs9660710
                                                                               0/0
                                                                                      0/0
                                                                                              0/0
                                                                                                      0/0
                                                                                                              1/1
                                                                                                                      0/0
                                                                                                                              0/0
                                                                                                                                      0/0
                                                                                                                                              0/0
                                                                                                                                                      0/0
                                                                                                                                                             0/1
                                                                                                                                                                     0/0
                                                                                                                                                                              0/0
                                                                                                                                                                                     0/0
                                                                                                                                                                                             0/0
                                                                                                                                                                                                     0/1
                                                                                                                                                                                                             0/0
                                                                                                                                                                                                                     0/0
                                                                                                                                                                                                                             0/1
                                                                                                                                                                                                                                    0/0
                                                                                                                                                                                                                                            0/0
        1106473 rs4970420
                                                                       GT
                                                                               0/0
                                                                                      0/0
                                                                                              0/1
                                                                                                      0/0
                                                                                                              0/0
                                                                                                                      0/0
                                                                                                                              0/1
                                                                                                                                      0/0
                                                                                                                                              0/0
                                                                                                                                                      0/0
                                                                                                                                                             0/1
                                                                                                                                                                     0/1
                                                                                                                                                                             0/0
                                                                                                                                                                                     0/1
                                                                                                                                                                                             0/0
                                                                                                                                                                                                     0/0
                                                                                                                                                                                                             0/0
                                                                                                                                                                                                                     0/0
                                                                                                                                                                                                                             0/1
                                                                                                                                                                                                                                            0/0
        1119858 rs1320565
                                                                                      9/9
                                                                                              9/9
                                                                                                              9/9
                                                                                                                      9/9
                                                                                                                                      9/9
                                                                                                                                              9/9
                                                                                                                                                             9/1
                                                                                                                                                                     9/9
                                                                                                                                                                                     a/a
                                                                                                                                                                                             9/9
                                                                                                                                                                                                             9/9
                                                                                                                                                                                                                                    9/9
```

Variant Call Format

```
##fileformat=VCFv4.3
##fileDate=20230307
##source=PLINKv2.00
##contig=<ID=1>
##contig=<ID=2>
##contig=<ID=3>
##contig=<ID=4>
##contig=<ID=5>
##contig=<ID=6>
##contig=<ID=7>
##contig=<ID=8>
##contig=<ID=9>
##contig=<ID=10>
##contig=<ID=11>
##contig=<ID=11>
##contig=<ID=12>
##contig=<ID=13>
##contig=<ID=14>
##contig=<ID=15>
##contig=<ID=16>
##contig=<ID=17>
##contig=<ID=18>
##contig=<ID=19>
##contig=<ID=20>
##contig=<ID=21>
##contig=<ID=22>
##contig=<ID=Y>
##contig=<ID=X>
```

0/0 ref. homozygous0/1 heterozygous1/1 alt. homozygous./. missing info

##INFO=	<id=pr, n<="" th=""><th>umber=0,Type=Fl:</th><th>ag, Descri</th><th>ption="P</th><th>rovision</th><th>al refer</th><th>ence all</th><th>ele, ma</th><th>y not be</th><th>based or</th><th>n real r</th><th>eference</th><th>genome"</th><th>></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th></id=pr,>	umber=0,Type=Fl:	ag, Descri	ption="P	rovision	al refer	ence all	ele, ma	y not be	based or	n real r	eference	genome"	>															
##FORMA	#FORMAT= <id=gt, description="Genotype" number="1," type="String,"></id=gt,>																												
#CHROM	POS	ID REF	ALT	QUAL	FILTER	INFO	FORMAT	HGDP00521_HGDP00521			HGDP00533_HGDP00533			HGDP01361_HGDP01361			HGDP01372_HGDP01372			HGDP00667_HGDP00667			HGDP01066_HGDP01066			HGDP01078_HGDP01078			HGDP01>
		rs12562034	G	A			PR	GT	0/0	0/0	0/0	0/1	0/0	0/0	0/0	0/0	0/0	0/0	0/1	0/0	0/1	0/0	0/0	0/0	0/1	0/0	0/1	0/0	0/0 >
		rs3934834					PR	GT	0/0	0/1	0/0	0/0	0/0	0/1	0/0	0/1	0/1	0/0	0/1	0/0	0/0	0/0	0/1	0/1	0/0	0/0	0/1	0/0	0/0 >
1		rs9442372	G	A			PR	GT	0/0	0/1	1/1	1/1	0/0	0/1	0/1	0/1	1/1	0/0	1/1	0/0	0/0	0/0	1/1	0/1	1/1	0/0	0/1	0/1	0/0 >
		rs9442398	G	A			PR	GT	0/0	0/0	1/1	1/1	0/0	0/0	0/0	0/0	0/1	0/0	0/1	0/0	0/0	0/0	0/1	0/0	1/1	0/0	0/0	0/1	0/0 >
1		rs6687776	С	T			PR	GT	0/0	0/0	0/0	0/1	0/0	0/0	0/0	0/0	0/1	0/0	0/1	0/1	0/1	0/0	0/0	0/1	0/0	0/1	0/1	0/0	0/0 >
1		rs9651273	G	A			PR	GT	0/0	0/1	1/1	0/0	0/0	0/1	0/1	0/1	0/1	0/0	0/1	0/0	0/0	0/0	1/1	0/0	1/1	0/0	0/0	0/1	0/0 >
1		rs4970405	A	G			PR	GT	0/0	0/0	0/0	0/1	0/0	0/0	0/0	0/0	0/1	0/0	0/1	0/1	0/0	0/0	0/0	0/1	0/0	0/1	0/1	0/0	0/0 >
1		rs12726255	A	G			PR	GT	0/0	0/0	0/0	0/1	0/0	0/0	0/0	0/0	0/1	0/0	0/1	0/1	0/0	0/0	0/0	0/1	0/0	0/1	0/1	0/0	0/0 >
1		rs7540009	G	A			PR	GT	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0 >
1		rs11807848		С			PR	GT	0/0	0/1	1/1	0/1	0/0	0/1	0/1	0/1	0/1	0/1	1/1	0/1	0/0	0/0	1/1	0/1	0/1	0/1	0/1	0/0	0/0 >
1		rs9442373	A	С			PR	GT	0/0	0/1	1/1	0/1	0/1	0/1	0/1	0/1	0/1	0/1	1/1	0/1	0/0	0/0	1/1	1/1	0/1	0/1	0/1	0/0	0/0 >
1		rs2298217	C				PR	GT	0/0	0/0	0/0	0/1	0/1	0/0	0/0	0/0	0/1	0/0	0/1	0/1	0/0	0/0	0/0	0/1	0/0	0/1	0/1	0/0	0/0 >
1		rs12145826	G	Ā			PR	GT	0/1	0/1	0/0	0/0	0/0	0/0	0/1	0/1	0/0	0/1	0/0	0/0	0/0	0/0	0/1	0/1	0/0	0/0	0/0	0/0	0/0 >
1		rs4970357	A	C			PR	GT	0/0	0/0	0/0	0/0	:/:	0/0	0/0	0/0	0/0	0/0	0/1	0/0	0/0	0/0	0/0	0/1	0/0	0/0	0/1	0/0	0/0 >
1		rs9442380	C	Ţ			PR	GT	0/0	0/0	0/0	0/0	1/1	0/0	0/0	0/0	0/0	0/0	0/1	0/0	0/0	0/0	0/0	0/1	0/0	0/0	0/1	0/0	0/0 >
1		rs7553429	A	C			PR	GT	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/1	0/0	0/0	0/0	0/0 >
1		rs4970362	G	A			PR	GT	0/0	0/0	0/1	0/0	1/1	0/0	0/0	0/0	0/0	0/0	1/1	0/0	0/0	0/0	0/0	0/1	0/0	0/0	0/1	0/0	0/1 >
1		rs9660710	C	A			PR	GT	0/0	0/0	0/0	0/0	1/1	0/0	0/0	0/0	0/0	0/0	0/1	0/0	0/0	0/0	0/0	0/1	0/0	0/0	0/1	0/0	0/0 >
1		rs4970420	G	A			PR	GT	0/0	0/0	0/1	0/0	0/0	0/0	0/1	0/0	0/0	0/0	0/1	0/1	0/0	0/1	0/0	0/0	0/0	0/0	0/1	1/1	0/0 >
1	1119858	rs1320565	c	Т			PR	RT	a/a	a/a	a/a	a/a	a/a	a/a	a/a	a/a	a/a	a/a	a/1	a/a	A/1	a/a	a/a	a/a	a/a	a/a	A/1	a/a	a/a 🔊

BED/BIM/FAM

BED

It is a BINARY file format.

0
1
2
0
1
1
0
1
1
2
1
2
1
2
2
1
2

BIM

Basically a tsv file for variants.

- 1. Chromosome code (either an integer, or 'X'/'Y'/'XY'/'MT'; '0' indicates unknown) or name
- 2. Variant identifier
- 3. Position in morgans or centimorgans (safe to use dummy value of '0')
- 4. Base-pair coordinate (1-based; limited to 2³¹-2)
- 5. Allele 1 (corresponding to clear bits in .bed; usually minor)
- 6. Allele 2 (corresponding to set bits in .bed; usually major)

<u>FAM</u>

Basically a tsv file for individual info.

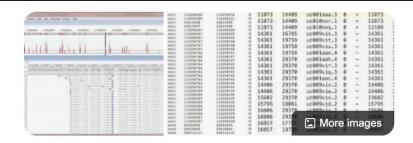
- 1. Family ID ('FID')
- 2. Within-family ID ('IID'; cannot be '0')
- 3. Within-family ID of father ('0' if father isn't in dataset)
- 4. Within-family ID of mother ('0' if mother isn't in dataset)
- 5. Sex code ('1' = male, '2' = female, '0' = unknown)
- 6. Phenotype value ('1' = control, '2' = case, '-9'/'0'/non-numeric = missing data if case/control)

Beware!



BED files

A BED file (. bed) is a tab-delimited text file that defines a feature track. It can have any file extension, but . bed is recommended. The BED file format is described on the UCSC Genome Bioinformatics web site: http://genome.ucsc.edu/FAQ/FAQformat.





Broad Institute

https://software.broadinstitute.org > software > igv > B...

BED | Integrative Genomics Viewer - Broad Institute

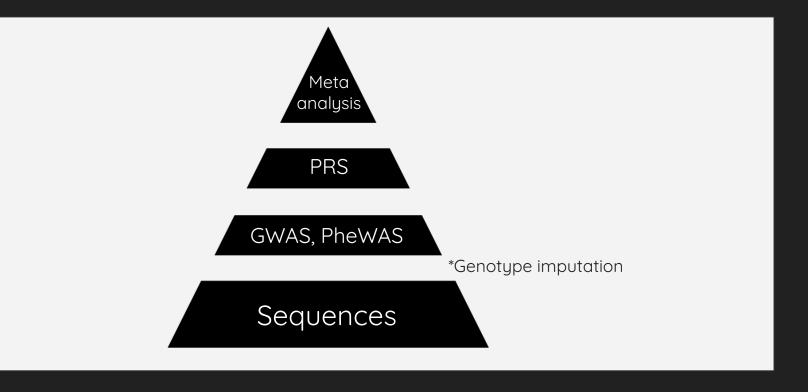


Exercise 1

Set-up and data exploration.

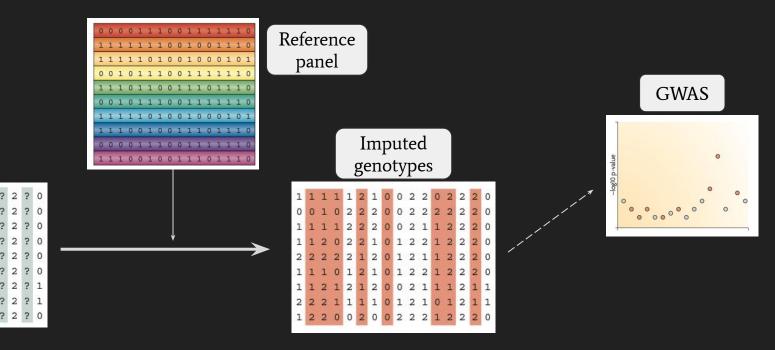
Genotype imputation

Basic structure

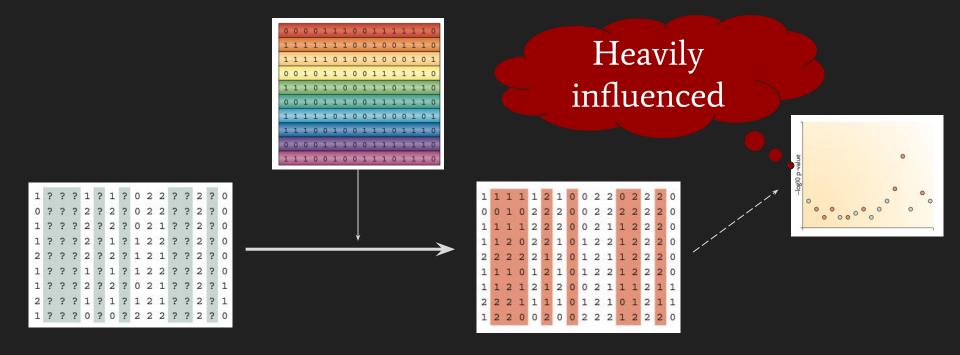


How does "genotype imputation" work?

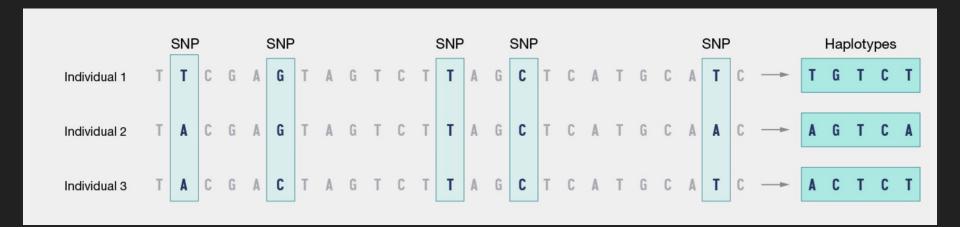
GSA



How does "genotype imputation" work?



Reference panel



Reference panel

Ancestry

General reference panels available at the moment:

- HapMap3
- 1000Genomes

Reference panel

Ancestry

Populations

The following population samples were studied:

ASW	African ancestry in Southwest USA
-----	-----------------------------------

CEU Utah residents with Northern and Western European ancestry from the CEPH collection

CHB Han Chinese in Beijing, China

CHD Chinese in Metropolitan Denver, Colorado

GIH Gujarati Indians in Houston, Texas

JPT Japanese in Tokyo, Japan

LWK Luhya in Webuye, Kenya

MXL Mexican ancestry in Los Angeles, California

MKK Maasai in Kinyawa, Kenya

TSI Toscani in Italia

YRI Yoruba in Ibadan, Nigeria

From theory to practice

Data preparation

Tools: plink2

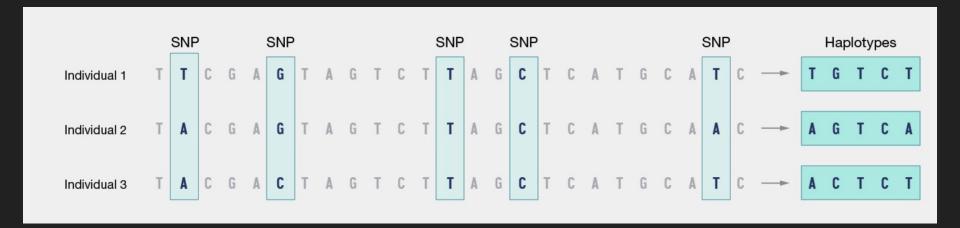
- 1. Correct input strand
- 2. Sort bcf
- 3. SNP quality control (maf, geno)
- 4. Individual quality control (mind)
- 5. Chromosome split
- 6. Remove duplicate variants

Exercise 2

Data preparation - Some basic filtering.

Phasing / Haplotype estimation

Shapeit4



State-of-the-art

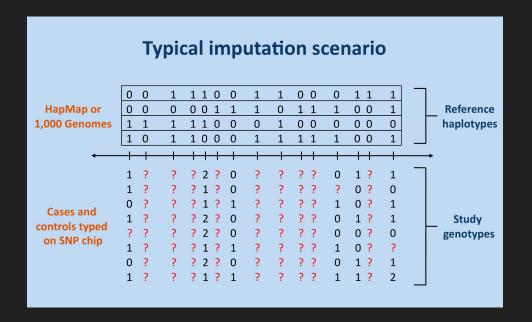
Imputation:

- 1. Impute5
- 2. Minimac4
- 3. BEAGLE

State-of-the-art

Imputation:

- 1. Impute5
- 2. Minimac4
- 3. BEAGLE



Exercise 3

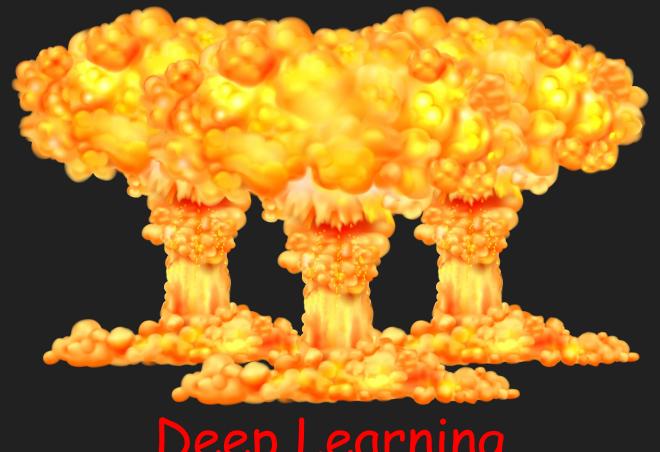
Let's impute some genotypes.

Post-imputation

Quality control of imputed SNPs:

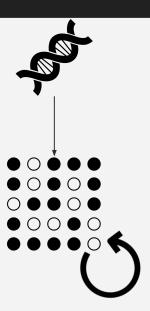
INFO SCORE -> range [0, 1]

Higher is better.



Deep Learning

Imputator



Rapid, Reference-Free human genotype imputation with denoising autoencoders

Raquel Dias^{1,2,3}, Doug Evans^{1,2}, Shang-Fu Chen^{1,2}, Kai-Yu Chen^{1,2}, Salvatore Loguercio^{1,2}, Leslie Chan^{1,2}, Ali Torkamani^{1,2}*

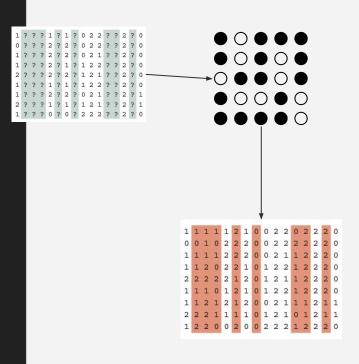
¹Scripps Research Translational Institute, Scripps Research Institute, La Jolla, United States; ²Department of Integrative Structural and Computational Biology, Scripps Research, La Jolla, United States; ³Department of Microbiology and Cell Science, University of Florida, Gainesville, United States

Abstract Genotype imputation is a foundational tool for population genetics. Standard statistical imputation approaches rely on the co-location of large whole-genome sequencing-based reference panels, powerful computing environments, and potentially sensitive genetic study data. This results in computational resource and privacy-risk barriers to access to cutting-edge imputation techniques. Moreover, the accuracy of current statistical approaches is known to degrade in regions of low and complex linkage disequilibrium. Artificial neural network-based imputation approaches may overcome these limitations by encoding complex genotype relationships in easily portable inference models. Here, we demonstrate an autoencoder-based approach for genotype imputation, using a large, commonly used reference panel, and spanning the entirety of human chromosome 22. Our autoencoder-based genotype imputation strategy achieved superior imputation accuracy across the allele-frequency spectrum and across genomes of diverse ancestry, while delivering at least fourfold faster inference run time relative to standard imputation tools.

Editor's evaluation

The paper describes a novel neural-network-based strategy for imputing unmeasured genotypes, which is a standard part of most association testing pipelines. The method is computationally intensive to train, but once training is complete the imputation is fast and accurate and does not require further access to a reference panel. It has the potential to be a practically-appealing alternative to existing methods. although further work (eg training of models) is required before this new approach can be applied genome-wide.

Imputator



Rapid, Reference-Free human genotype imputation with denoising autoencoders

Raquel Dias^{1,2,3}, Doug Evans^{1,2}, Shang-Fu Chen^{1,2}, Kai-Yu Chen^{1,2}, Salvatore Loguercio^{1,2}, Leslie Chan^{1,2}, Ali Torkamani^{1,2}*

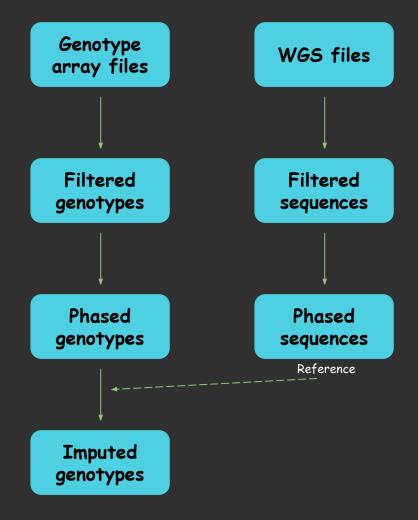
¹Scripps Research Translational Institute, Scripps Research Institute, La Jolla, United States; ²Department of Integrative Structural and Computational Biology, Scripps Research, La Jolla, United States; ³Department of Microbiology and Cell Science, University of Florida, Gainesville, United States

Abstract Genotype imputation is a foundational tool for population genetics. Standard statistical imputation approaches rely on the co-location of large whole-genome sequencing-based reference panels, powerful computing environments, and potentially sensitive genetic study data. This results in computational resource and privacy-risk barriers to access to cutting-edge imputation techniques. Moreover, the accuracy of current statistical approaches is known to degrade in regions of low and complex linkage disequilibrium. Artificial neural network-based imputation approaches may overcome these limitations by encoding complex genotype relationships in easily portable inference models. Here, we demonstrate an autoencoder-based approach for genotype imputation, using a large, commonly used reference panel, and spanning the entirety of human chromosome 22. Our autoencoder-based genotype imputation strategy achieved superior imputation accuracy across the allele-frequency spectrum and across genomes of diverse ancestry, while delivering at least fourfold faster inference run time relative to standard imputation tools.

Editor's evaluation

The paper describes a novel neural-network-based strategy for imputing unmeasured genotypes, which is a standard part of most association testing pipelines. The method is computationally intensive to train, but once training is complete the imputation is fast and accurate and does not require further access to a reference panel. It has the potential to be a practically-appealing alternative to existing methods. although further work (eg training of models) is required before this new approach can be applied genome-wide.

Pipeline overview



Thank you