The Netherlands Twin Register Axiom Biobank platform for GWAS

J.J. Hottenga[1,2], E. Ehli[3], A. Abdellaoui[1], I. Fedko[1], G. Willemsen[1], E.J.C.N. de Geus[1,2], C. Grieser[4], G.E. Davies[3], D.I. Boomsma[1]

1) The Netherlands Twin Register, VU University, Amsterdam, NL; **2)** EMGO institute for Health and Care Research, VU University and VU University Medical Center, Amsterdam, NL; **3)** Avera Institute for Human Genetics, Sioux Falls, SD, USA; **4)** Affymetrix, Santa Clara, CA, USA.

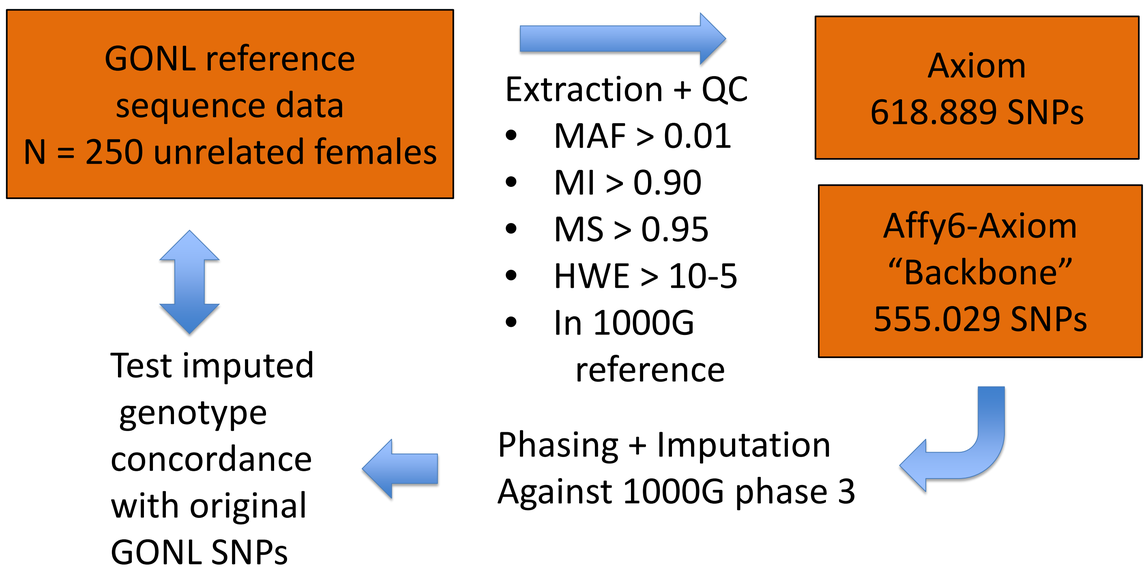
On the NTR Axiom

Function	**#SNPs**
Imputation Axiom + Affymetrix 6 SNPs	566.515
Improved coverage X, Y and Mitochondrial	2.187
Neurological, developmental, twinning and psychiatric	50.153
Cardio vascular, metabolomics, muscle - and lung function	16.714
Immune function and disorders, HLA, KIR	9.152
Pharmacogenomics	2.037
Cancer common variants	343
GWAS catalog loci for many phenotypes	11.036
EQTLs	17.115

Usage

- Genotyping many known GWAS loci
- Common variant risk profiling
- Good imputation coverage
- Backwards compatible with Affymetrix 6.0
- Zygosity and IBD determination

Coverage Methods



Results

Figure 1 : ~10M polymorphic GONL SNPs re-imputed with the 1000G reference (MAF>0.005). The graph represents the % of SNPs that were covered well, ok or bad based on the concordance between the original and re-imputed genotypes.

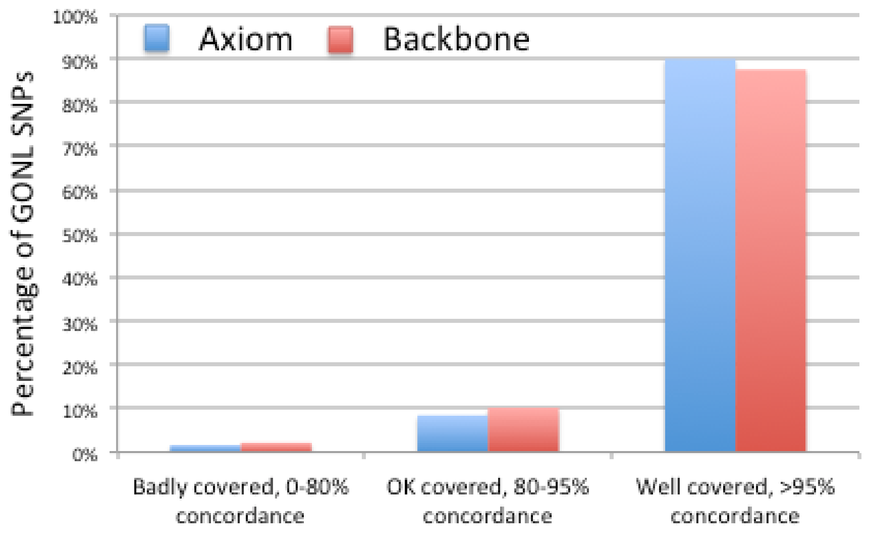


Figure 2: The median R^2 of the imputed SNPs in the full 1000G imputation ~41M SNPs, or only overlapping SNPs of GONL and 1000G ~13M SNPs (MAF>0 in 250 females)

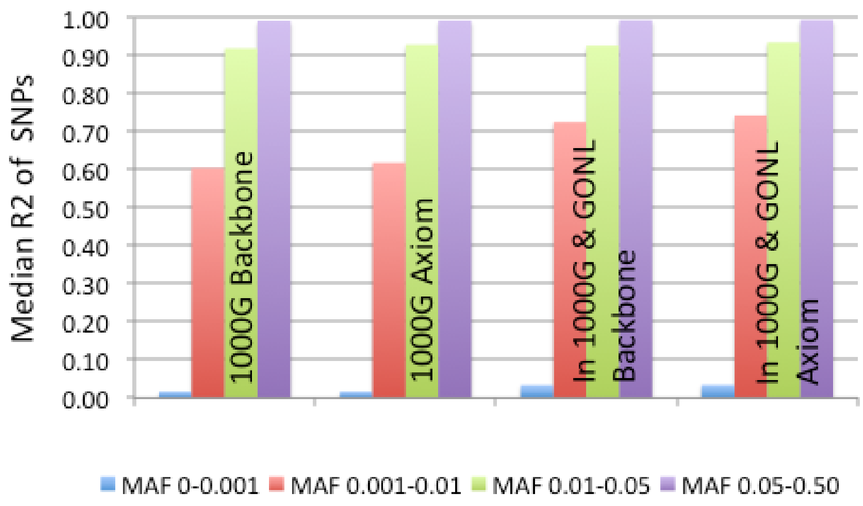
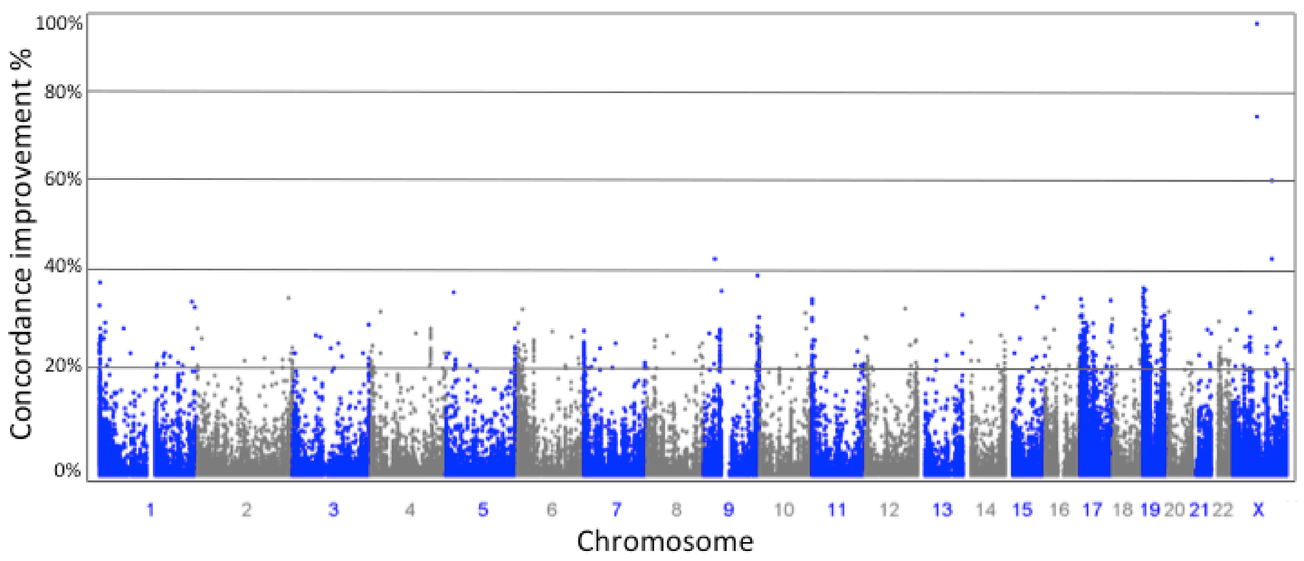


Figure 3: Locations of SNPs (N=115939) with a concordance difference >3 sd (0.05) of the mean (0.002) for Backbone vs. Axiom imputation.



Conclusions

The NTR axiom design captures the Dutch genetic variation very well using a standard 1000G imputation.
It allows rapid genotyping of a large number of highly relevant SNPs in collected DNA samples. When combined with existing Affy 6 genotyping, there is no substantial loss of information genome wide, however SNPs all over the genome are going to be imputed with less accuracy as compared to an Axiom only data set.