

# Expression estimation and eQTL mapping of HLA genes

Meyer Lab<sup>1</sup>    Dermitzakis Lab<sup>2</sup>

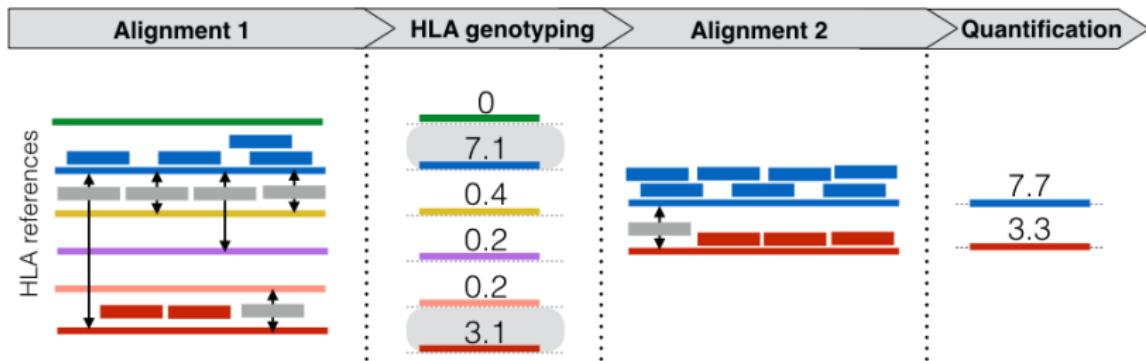
<sup>1</sup>University of São Paulo

<sup>2</sup>University of Geneva

## Motivation

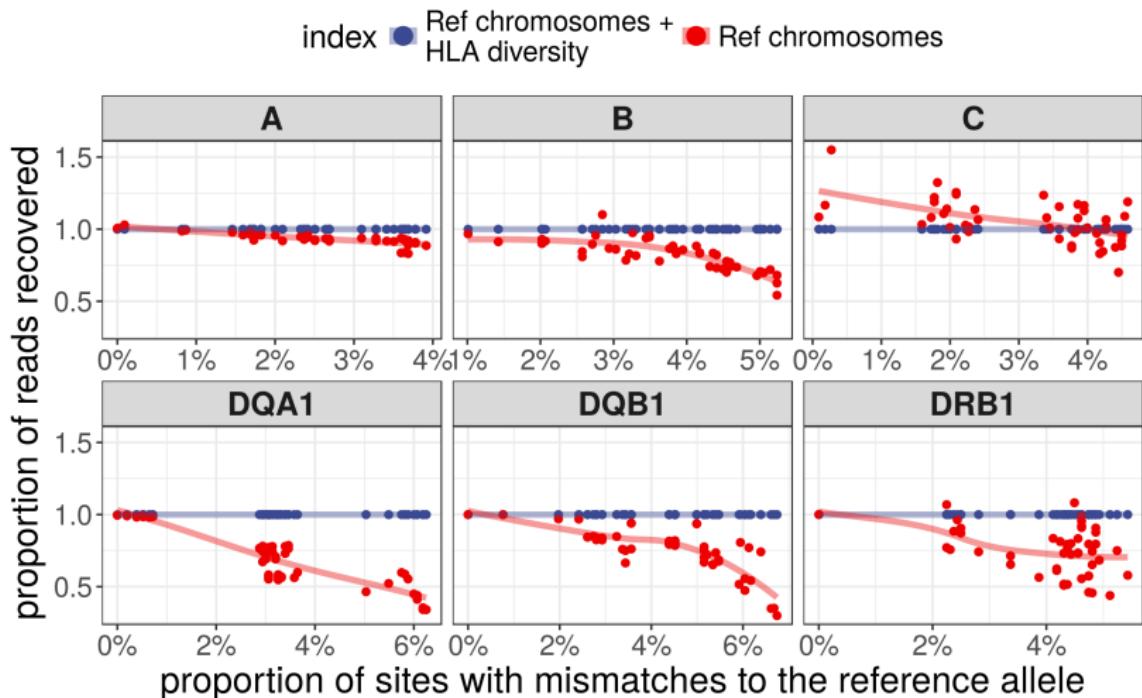
- ▶ Association to cancer phenotypes
- ▶ Association to HIV/AIDS progression
- ▶ Transplantation
- ▶ Many datasets of genome-wide data are available, but HLA estimates may be inaccurate
- ▶ Better eQTL mapping
- ▶ HLA allele-level estimates
- ▶ Understand how the regulatory architecture is related to HLA lineages

# Strategy



# Simulation: Indices covering HLA diversity recover more reads

Typing accuracy: 100%



## Quality assessment

Gene	not aligned	aligned to different ref	false positive reads
HLA-A	0	1.66	0
HLA-B	0	0.2	0
HLA-C	0	0	0
HLA-DQA1	0	0.4	0
HLA-DQB1	0	0	0
HLA-DRB1	0	0.4	0

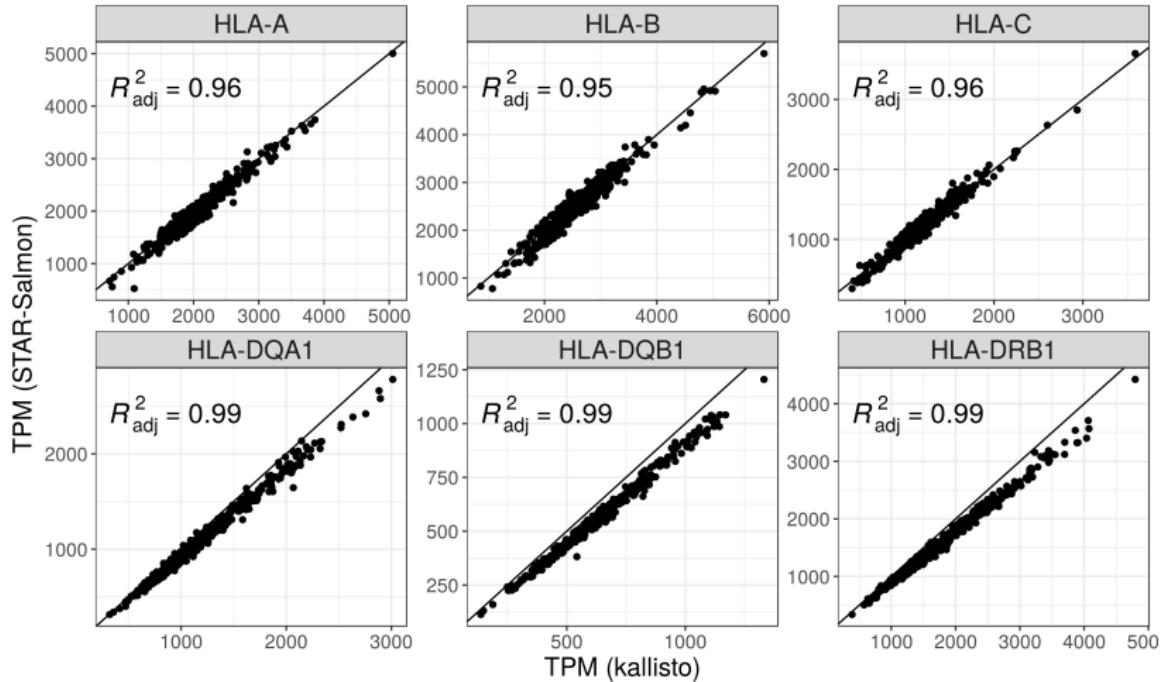
# Geuvadis Data

## Typing accuracy

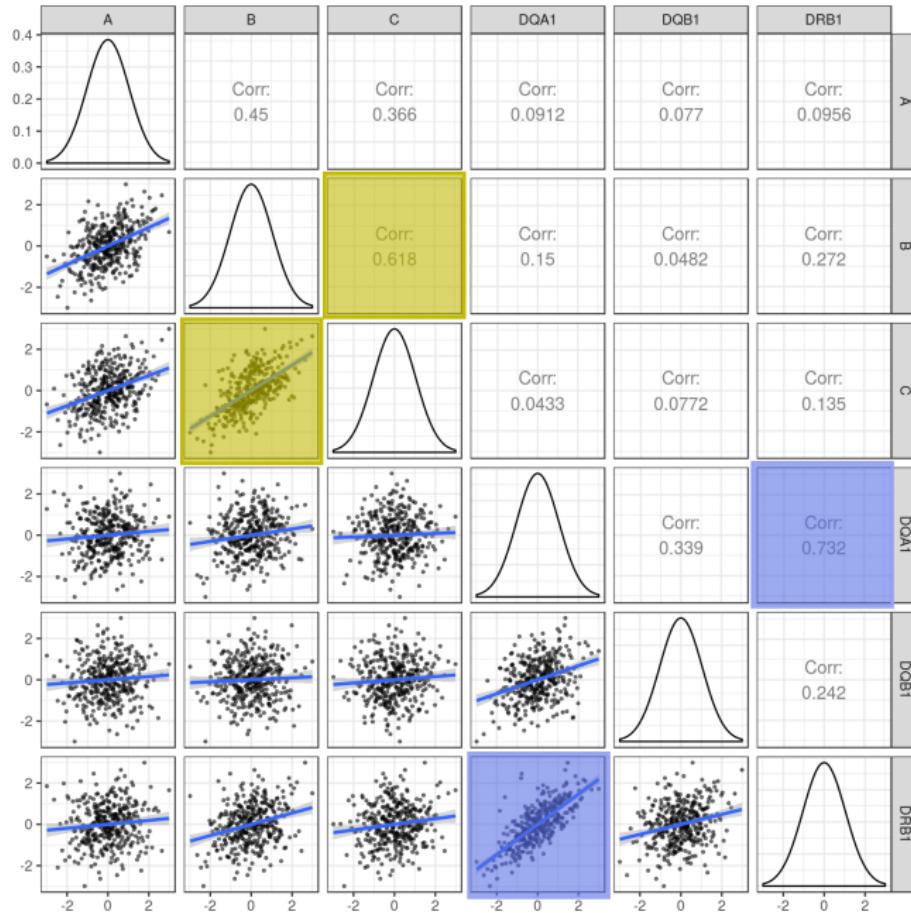
Gene	accuracy
HLA-A	97%
HLA-B	98%
HLA-C	96%
HLA-DQB1	95%
HLA-DRB1	98%

\*Gold standard: 1000 Genomes individuals HLA-typed by Gourraud et al. (2014). PLoS ONE 9(7): e97282.

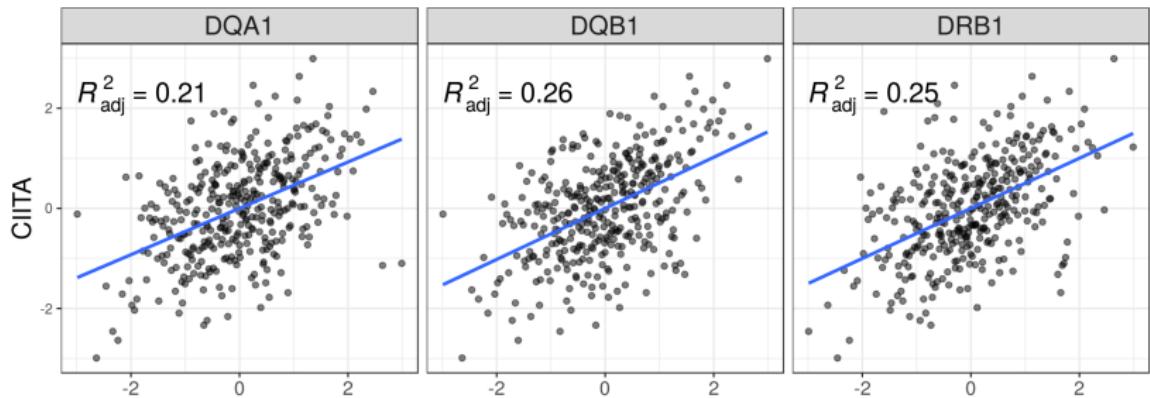
# Aligner choice does not influence expression estimate



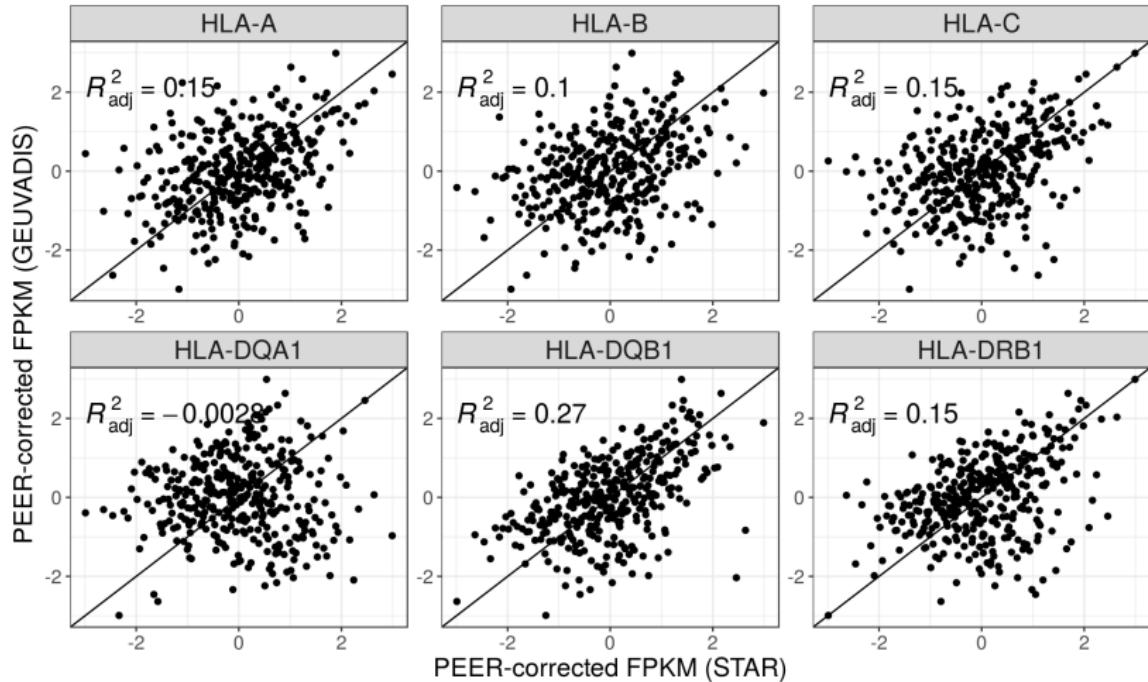
# Validation 1: Correlation of expression



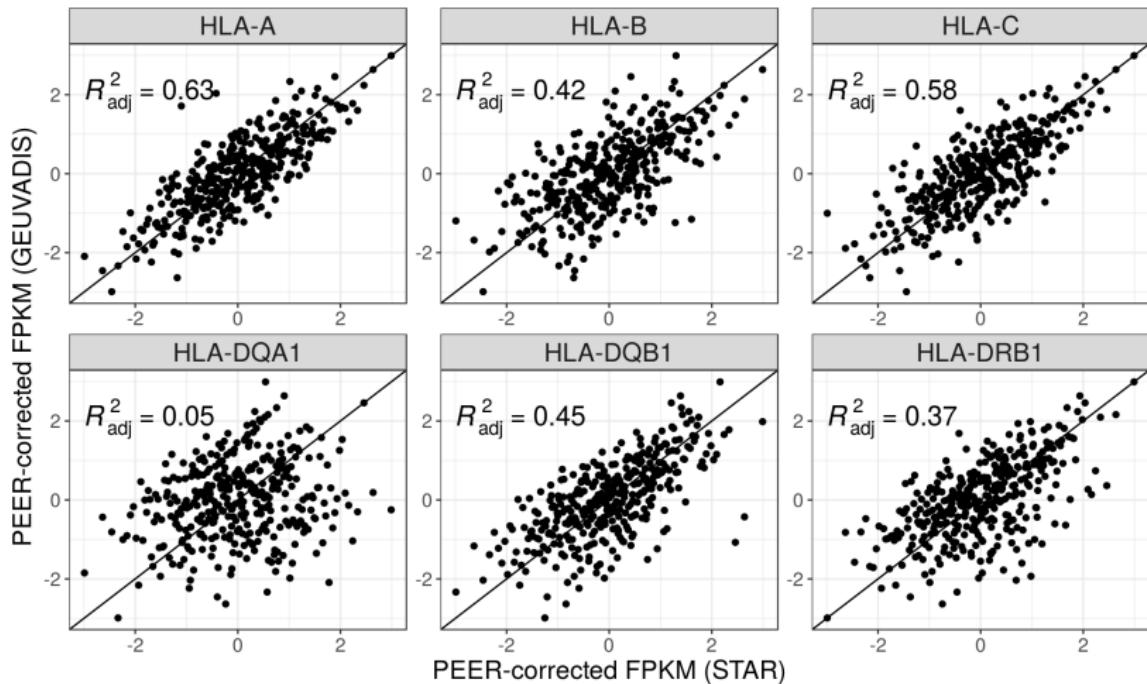
## Validation 2: Class II genes and their transactivator



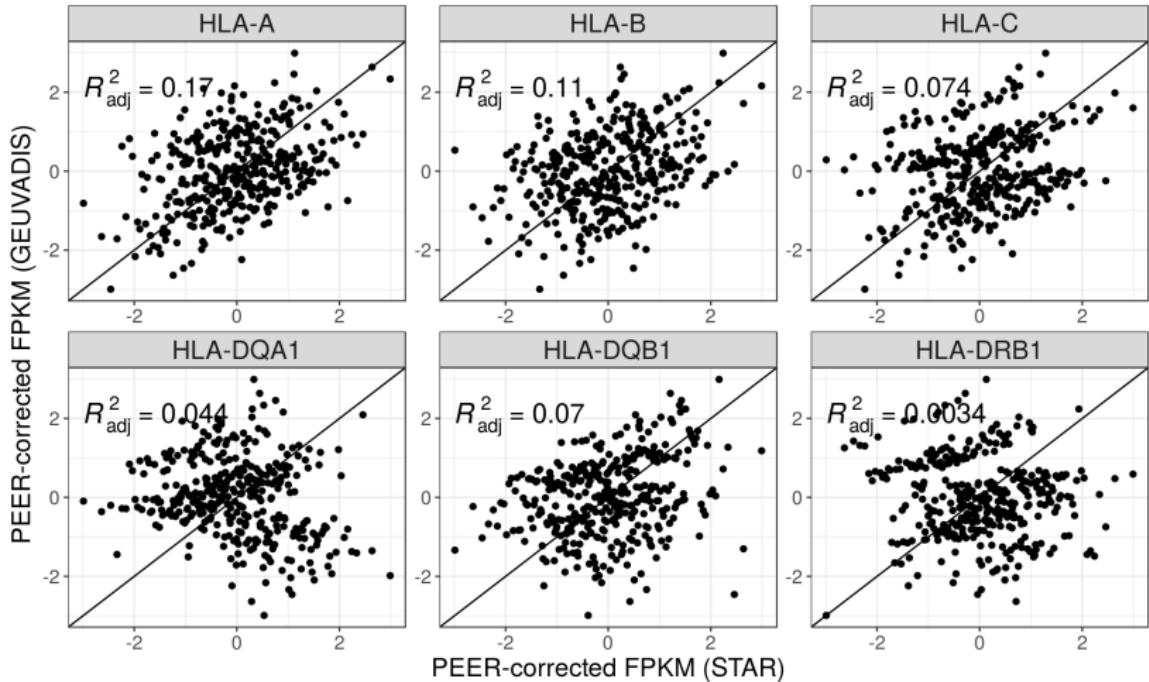
# Correlation with Geuvadis published estimates



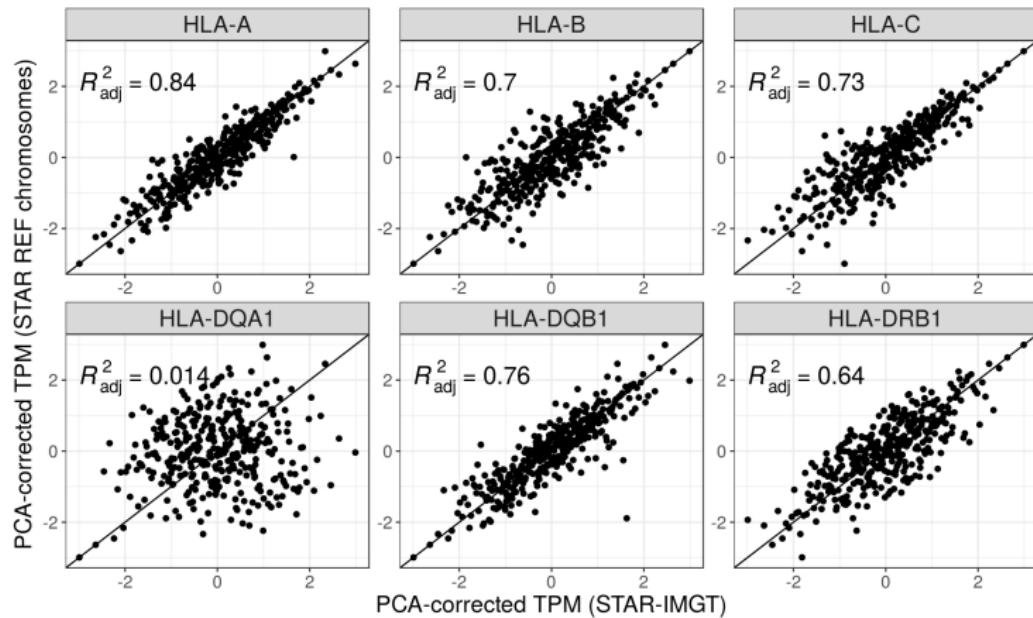
# Taking the uncorrected Geuvadis data and running myself the correction



# New quantifications

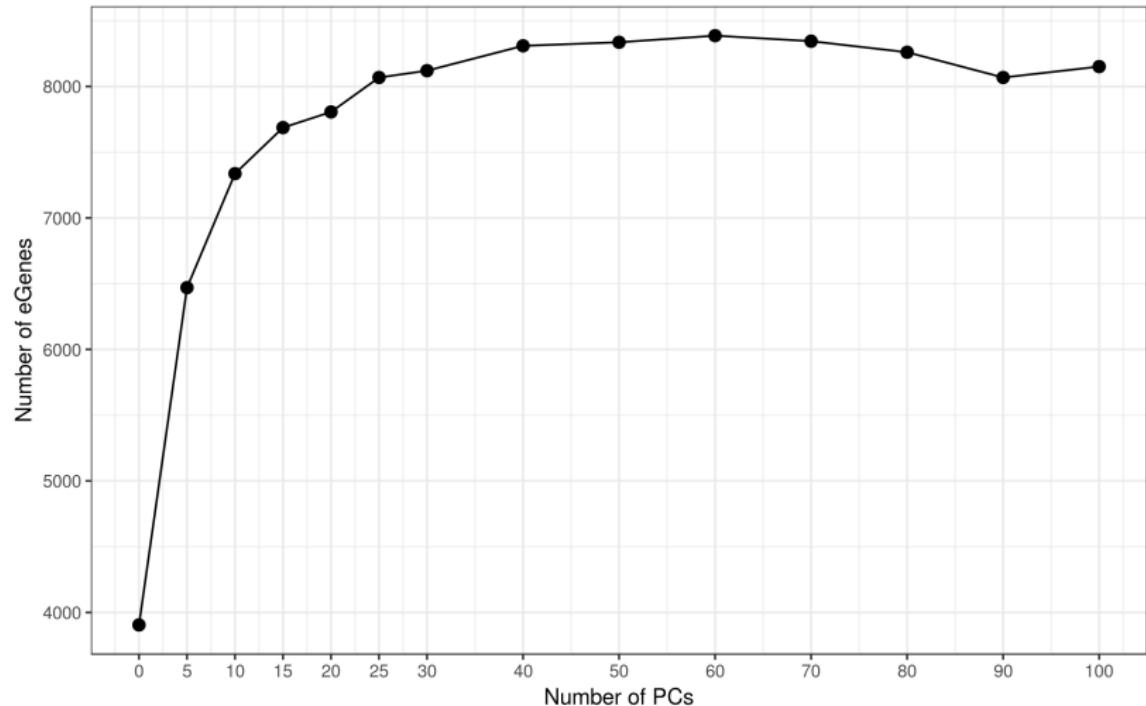


# Mapping to REF chromosomes may be enough for gene-level estimates

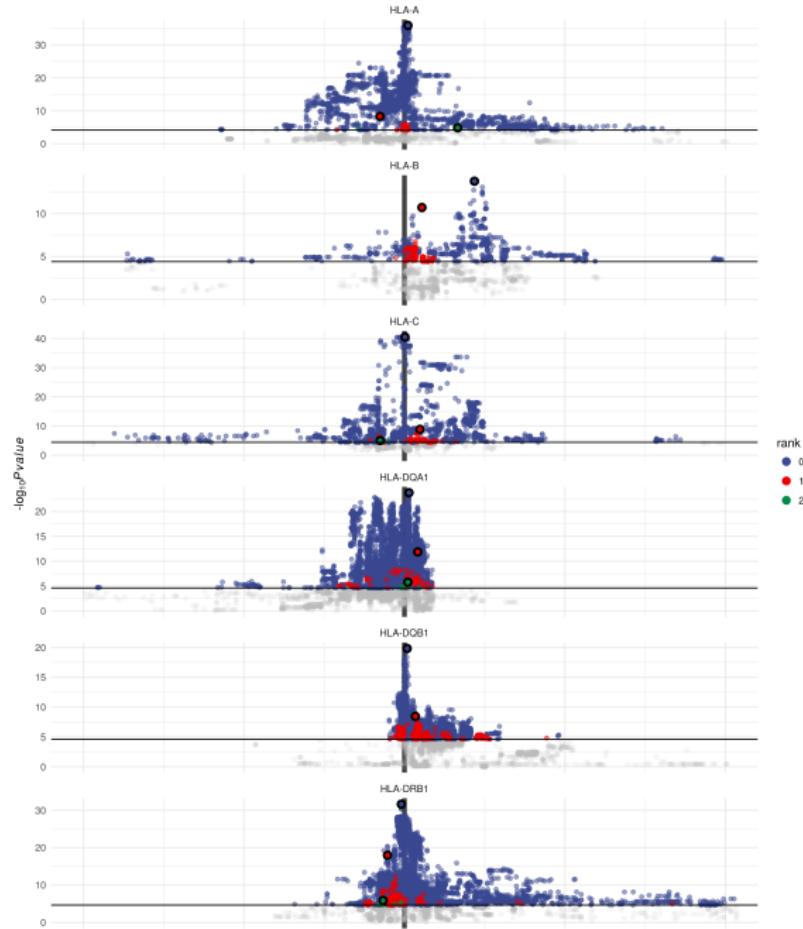


eQTLs

# Number of eGenes



# eQTLs at HLA genes



rank

0

1

2

# RTC with Geuvadis eQTLs

Gene	rank	variant	geuvadis gene	geuvadis variant	RTC
HLA-C	0	rs41561715	HLA-C	rs115899777	0.95
HLA-DQA1	0	rs373884000	HLA-DRB1	rs116405062	0.97
HLA-DQA1	1	rs9274722	HLA-DQA1/HLA-DQB1	rs9274660	0.92
HLA-DQB1	0	rs1770	HLA-DQA1/HLA-DQB1	rs9274660	0.98
HLA-DRB1	0	rs9270698	HLA-DQA1/HLA-DQB1	rs9274660	0.92
HLA-DRB1	2	rs145607970	HLA-DQA1/HLA-DQB1	rs9274660	0.93

# How allele-level expression is estimated in the HLA field

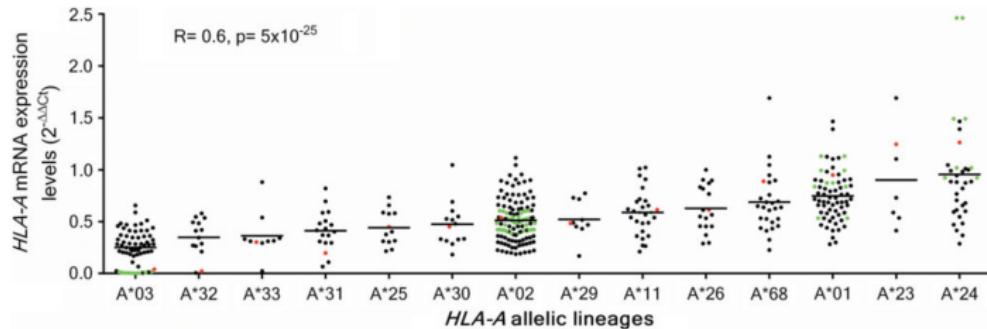
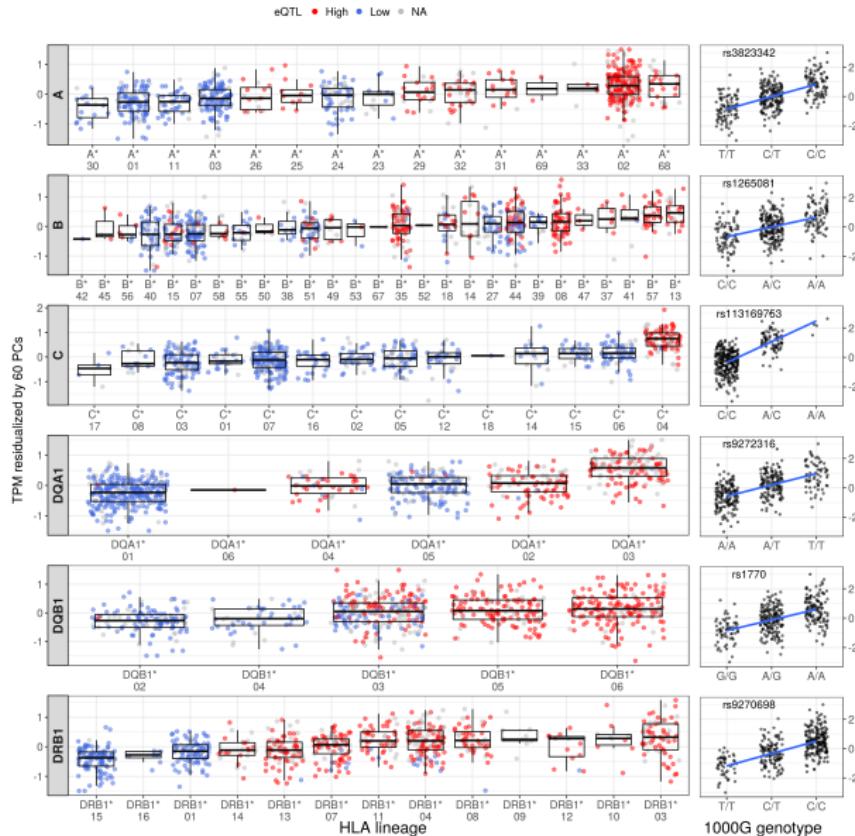


Figure 1. HLA-A gradient of expression across allelic lineages present in EAs. HLA-A mRNA expression level in 216 healthy donors was measured using qPCR. The average expression level of HLA-A was normalized to  $\beta$ 2M mRNA using the  $2^{-\Delta\text{CT}}$  method. The relative expression value is plotted twice for each donor (i.e. once for each HLA-A lineage present). Expression levels correlate significantly with HLA-A lineage and are continuously distributed ( $R = 0.6$ ,  $P = 5 \times 10^{-25}$ , ANOVA). Allelic lineages with  $\geq 5$  individuals are shown. The horizontal line indicates the raw average expression for the specific lineage. HLA-A homozygous individuals are marked in green dots. Red dots represent linear regression estimates of average expression levels for each HLA-A allelic lineage (estimated as a diploid homozygote).

Ramsuram et al. Human Molecular Genetics, 2015, 24 (15).

# Relationship between eQTLs and HLA alleles



# expression of alleles from different genes on the same vs on different haplotypes

