

VU Research Portal

(Epi) genetics and twins

van Dongen, J.

2015

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

van Dongen, J. (2015). *(Epi) genetics and twins*. [PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam].

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Appendix 4

Genome-wide DNA Methylation data NTR

Subjects and samples

The subjects in this study participated in the Netherlands Twin Register (NTR)^{1,2} biobank project³. Most subjects were twin pairs. A small number of subjects were family members of twins who were enrolled in the Genome of the Netherlands project (GONL⁴; which included parent-offspring trios where the offspring were twins or siblings of twins), or in the Complete Genomics project (CG; which included a few spouses of twins). Venous blood samples were drawn in the morning after an overnight fast, and separate EDTA tubes were collected for isolation of DNA and assessment of haematological profiles. Blood sampling and buccal sample collection procedures were described in detail previously³.

For the genome-wide methylation study 3264 blood samples from 3221 subjects were assessed: For 3178 subjects, a single blood sample was assessed and two blood samples were assessed for 43 subjects who participated twice in the NTR biobank (Biobank 1 and Biobank 2). After quality control (QC) of the methylation data, genome-wide methylation data were available for 3089 peripheral blood samples (mean age at blood draw=36.9, SD 12.9, range 17-79) from 3057 subjects (65.7% female; 34.3% male; GONL=331, CG=16) of which 32 subjects had methylation data for two samples (mean follow-up time=5 years, range=2.8-7.0). Subjects came from 1601 NTR families (Table 1). The post-QC whole blood DNA methylation dataset included 1250 complete twin pairs (MZ:803 (male= 239, female= 564), DZ:447 (male= 97, female=193, opposite-sex=157).

Table 1: Composition of families included in the whole blood genome-wide DNA methylation dataset (after DNA methylation Quality Control).

Description	Family size*					Total N families
	1	2	3	4	5	
1 twin	268	0	0	0	0	268
2 twins or 2 sibs	0	1217	0	0	0	1217
1 parent	7	0	0	0	0	7
1 parent + 1 child	0	2	0	0	0	2
1 parent + 2 children	0	0	3	0	0	3
2 parents, no children	0	8	0	0	0	8
2 parents, 1 child	0	0	64	0	0	64
2 parents, 2 children	0	0	0	25	0	25
2 parents, 3 children	0	0	0	0	4	4
1 spouse	1	0	0	0	0	1
1 spouse (of twin) + 2 twins	0	0	2	0	0	2
Total N families	276	1227	69	25	4	1601
Total N subjects	276	2454	207	100	20	3057

* Number of individuals per family

Genotype data used for imputation against the GONL reference panel for the BIOS consortium

Of the 3221 subjects for whom peripheral blood methylation samples were assessed with the illumina 450k array, 349 subjects were part of GONL, of whom 347 subjects were successfully sequenced (two subjects were excluded during quality control of the sequence data). Of the remaining 2872 subjects who were not part of the GONL study, 2665 subjects had been previously genotyped or had a MZ co-twin who had been genotyped one or multiple times on any of the following genotype arrays: Affymetrix6, Affymetrix 5, and Illumina660. The SNP data from the Affymetrix6, Affymetrix 5, and Illumina660 array were used for imputation against the GONL reference panel by the Biobank-based Integrative Omics Study (BIOS) consortium, in which a subset of the NTR DNA methylation samples takes part. Prior to imputation, one set of genotypes was selected (the one with the best quality) for MZ twins if both twins were genotyped and for individuals who had been genotyped on multiple platforms. Because the majority of subjects had been genotyped with Affymetrix 6, while only a small proportion had been genotyped with Affymetrix 5 or Illumina660, the individuals genotyped on Affymetrix 6 were separately imputed. The Affymetrix5 and Illumina660 data plus a small subset of subjects for whom only raw affymetrix6 genotypes were available that had not yet been cleaned (in December 2013) were merged and this combined dataset was imputed separately, after selecting the overlapping SNPs from Affymetrix5 and Illumina660 and (raw) Affymetrix 6 data. The following number of subjects were included in each set (including 1 subject from MZ pairs): affymetrix6; N=1621 subjects, merged SNP data set from other chips; N= 249 subjects. In total, there were 795 subjects whose own genotypes were not imputed but who were part of an MZ twin pair of which the co-twin was genotyped on Affymetrix6, Affymetrix5 or Illumina660.

Reference List

1. Boomsma,D.I. *et al.* Netherlands Twin Register: from twins to twin families. *Twin Res Hum Genet* **9**, 849-857 (2006).
2. Willemsen,G. *et al.* The adult Netherlands twin register: twenty-five years of survey and biological data collection. *Twin. Res. Hum. Genet.* **16**, 271-281 (2013).
3. Willemsen,G. *et al.* The Netherlands Twin Register biobank: a resource for genetic epidemiological studies. *Twin Res Hum Genet* **13**, 231-245 (2010).
4. Boomsma,D.I. *et al.* The Genome of the Netherlands: design, and project goals. *Eur. J. Hum. Genet.* **22**, 221-227 (2014).