

## OMIKDATA I THL BIOBANK (18.3.2026)

I tabellerna visas antalet av olika typer av omikdata i THL Biobanken. För mer information kontakta THL Biobanken: biopankki (at) thl.fi.

### 1. Befolkningsundersökningar

Samlingens namn	Imputerade genotyper	Chip-genotyper	WES (vcf format)	WGS (vcf format)	Metabolomik (NMR) *	Övriga omikdata
FINRISK 1992	4930	5200	1220	530	N/A	N/A
FINRISK 1997	7030	7465	3150	1280	7600	N/A
FINRISK 2002	7150	7460	5150	1980	7920	N/A
FINRISK 2007	5200	5370	2420	670	6010	Proteomik: 15 (Olink)
FINRISK: DILGOM 2007 (återkallelse av deltagare från FINRISK 2007)	4700 ***	4730 ***	2040 ***	560 ***	4815	Transkriptomik: 515 Methylomik: 510 Telomerdata:4070
FINRISK: DILGOM 2014 (återkallelse av deltagare från FINRISK 2007)	1210 ***	1230 ***	N/A	N/A	1250 (1230 med 7 års uppföljning)	Transkriptomik: 325 (alla med 7 års uppföljning)
FINRISK 2012 **	5460	5800	1915	N/A	5520	Proteomik: 10 (Olink)
Hälsa 2000	6550	7000	4690	205	7140	Telomerdata: 7400
Hälsa 2011	4400 (varav 3 840 också i H2000)	4700 (varav 4000 också i H2000)	2500 (varav alla också i H2000)	210	4750 (3870 med 11 års uppföljning)	N/A
FinHälsa 2017	6100	6400	N/A	N/A	5290	Proteomik: 30 (Olink)

\* NMR-metabolomiken har analyserats med Nightingale Healths plattform

\*\* FINRISK 2012-antalet inkluderar FINRISK: Kuusamo hälsoundersökning 2011

\*\*\* DILGOM 2007 och 2014 är genotypade och sekvenserade från FINRISK 2007-provet

## 2. Övriga samlingar

Samlingens namn	Imputerade genotyper	Chip-genotyper	WES (vcf format)	WGS (vcf format)	Metabolomik (NMR) *	Övriga omikdata
SUPER-studien	8570	9020	8350	870	N/A	Proteomik: 125 (Olink) 10 (NULISA)
GeneRISK-studien	6970	7270	N/A	N/A	N/A	Proteomik: 50 (Olink) Lipidomik: 7260
THL Biobankens GWAS-kohort	152 400	153 000	33 500	5 800	37 300	Proteomik: 420 (Olink) 180 (NULISA) Lipidomik: 7250

\* NMR-metabolomiken har analyserats med Nightingale Healths plattform

### Data formats:

- Genomdata imputerade i FinnGen finns som VCF-filer (GRCh38/hg38).
- Chip-genotyper finns som Plink- eller VCF-filer (hg18/hg19/hg38).
- WES-data finns som VCF-filer (GRCh37/hg19 eller GRCh38/hg38).
- WGS-data finns som VCF-filer (GRCh38/hg38).
- Transkriptomikdata för DILGOM 2007 och 2014 använder NCBI36/hg18.

**THL Biobankens imputationsreferenspanel**, N=2827 (en del av SISu v4 referenspanel), finns i vcf format (GRCh38/hg38) och får användas endast för imputering och kan inte kopplas till fenotypdata.