Supplementary Materials:

Figure. S1: Integrating longitudinal gene expression data from BABYDIET and DIPP cohorts for differential gene expression analyses. A) A schematic view of data integration. Most data were sampled before seroconversion in BABYDIET and after seroconversion in DIPP. Gene expression was also higher in BABYDIET compared to DIPP. Variance stabilised normalization was used for both cohorts. The cohort effect was removed using the limma R package.

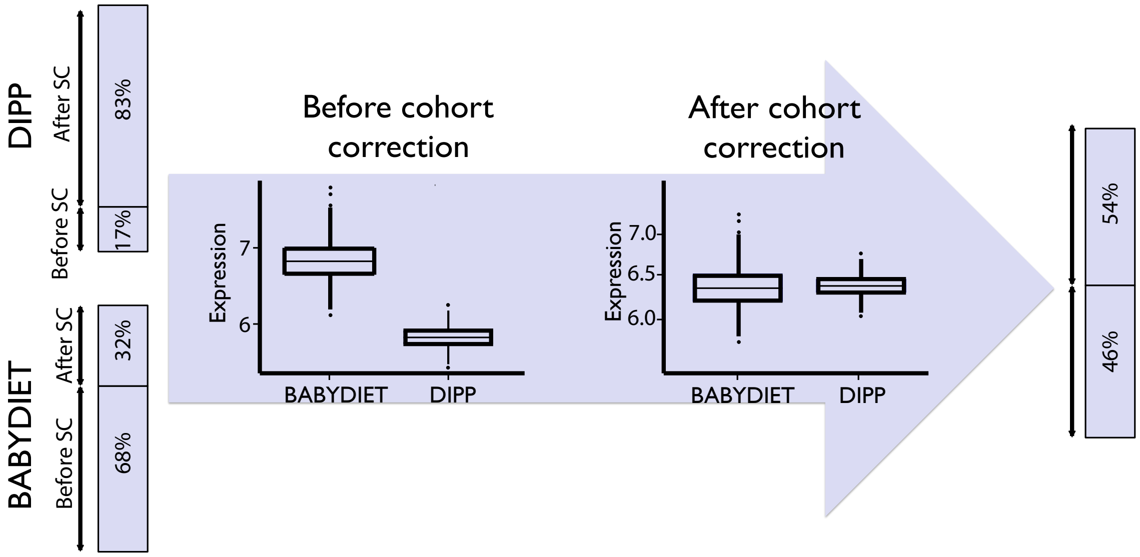


Figure. S2: Computational deconvolution of DE genes. A) Deconvoluted cell-specific proportion based on genomic analysis in seroconvertors and B) non seroconvertors from the DIPP cohort over time.

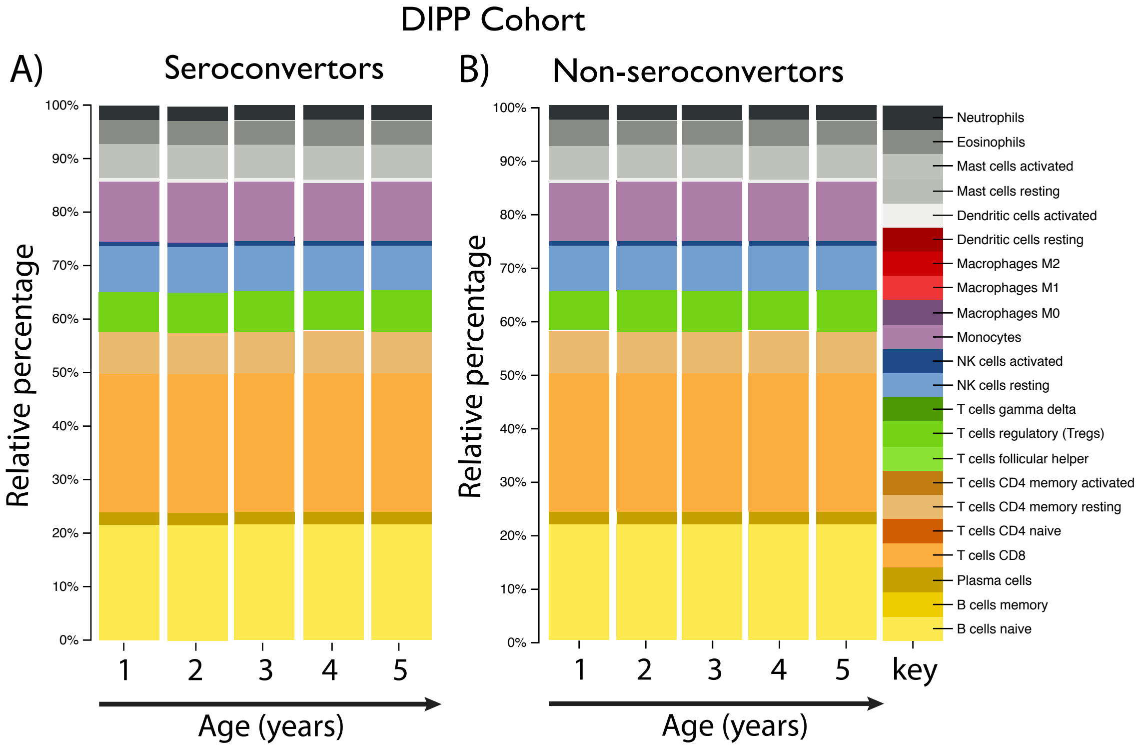


Figure. S3: Power analysis of at/near birth data. The relationship between Type-1 error and power is shown.

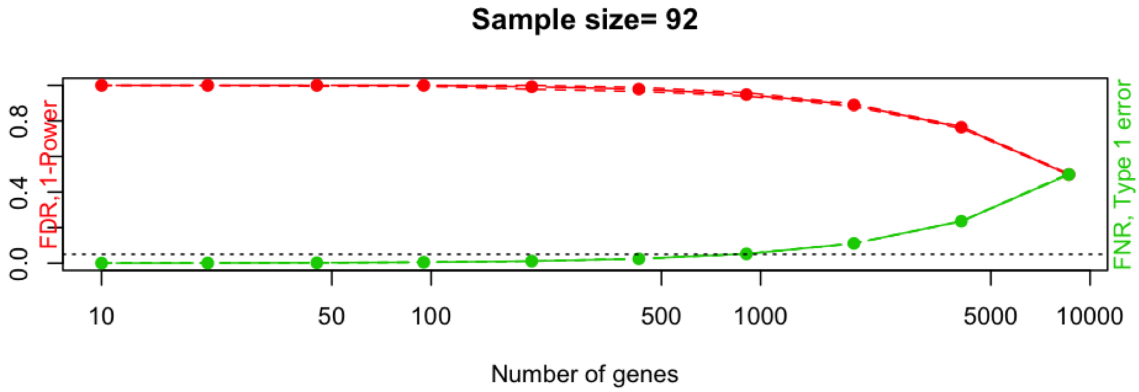


Figure. S4: Selection of informative features predicting seroconversion using random forest method as a pre-filter. A) 20 features pre-filtered for step-wise logistic regression method to predict seroconversion. Mean decrease in Gini Index was used to assess variable importance. Together with clinical variables (age, HLA score, gender), these features were then included in a step-wise logistic regression model.

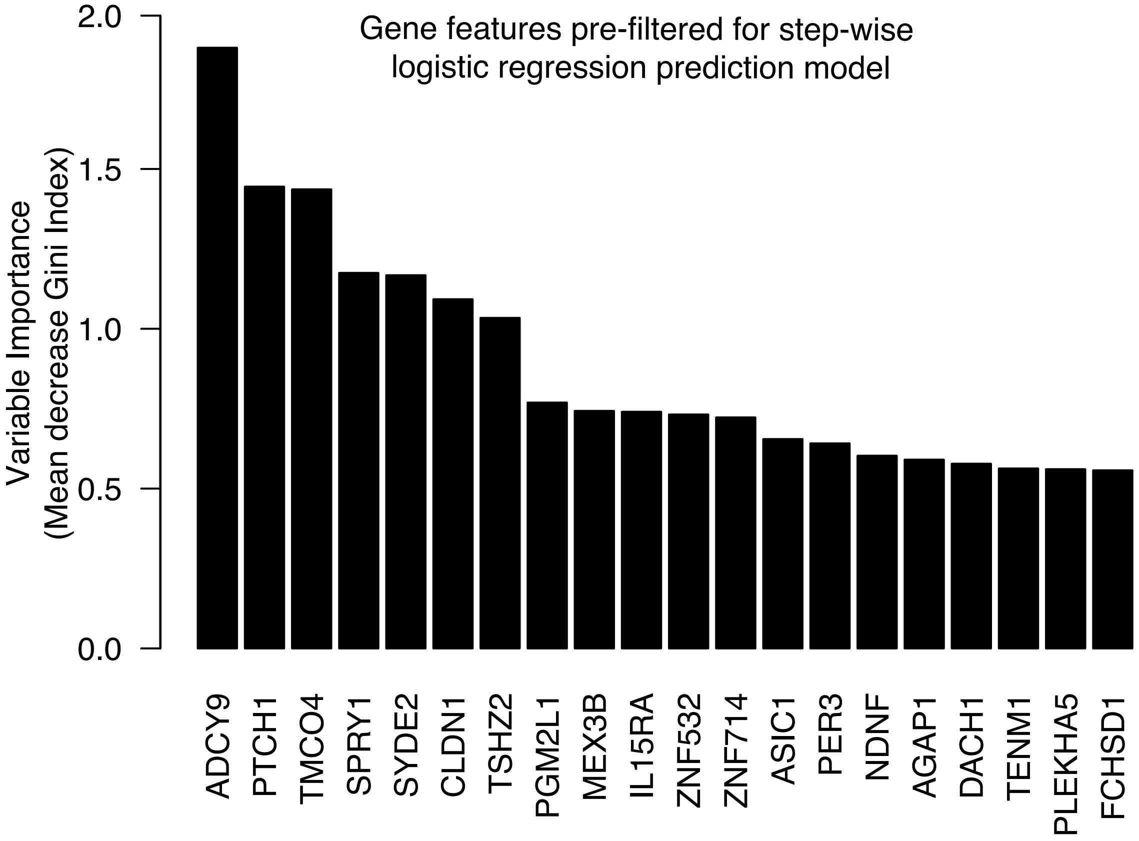
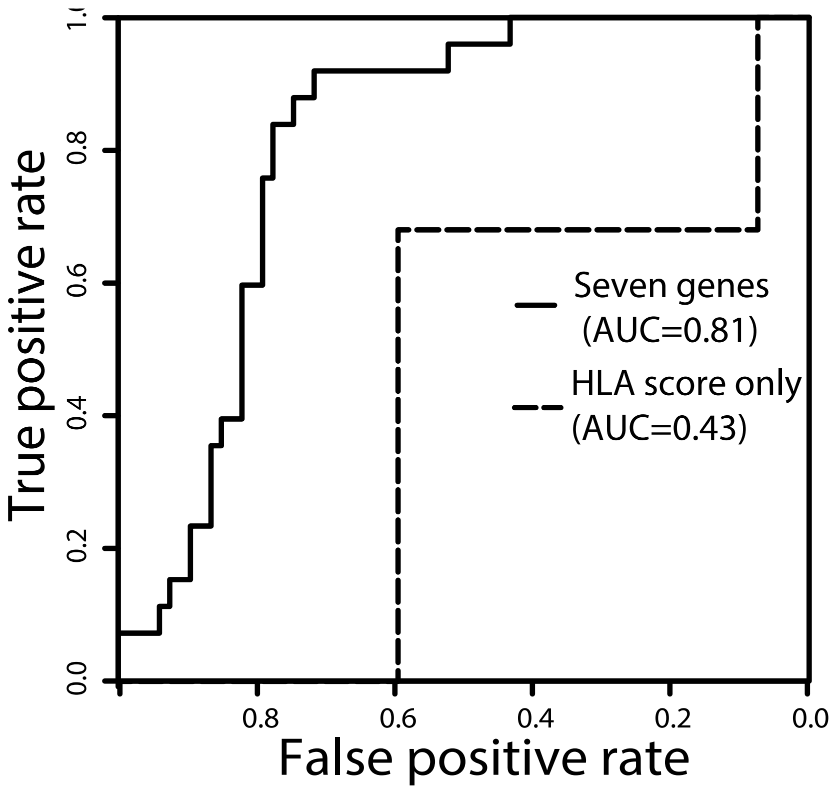


Figure. S5: Contribution of genes and HLA to seroconversion prediction. The continuous line ROC curve represents the prediction of seroconversion by using seven genes only. The dotted ROC curve represents the contribution of HLA-score only. AUC values for both curves are also shown.



Supplementary Tables

Table S1: Sensitivity analyses on differentially expressed statistically significant genes filtered a subset of 67 differentially expressed genes.

To remove any bias in our analysis that may result from differences in clinical characteristics between the seroconvertors and non-seroconvertors, we performed sensitivity analyses, where we randomly permuted samples containing 110 genes by matching subject number, gender, HLA-risk and microarray data points and assessed their differential expression. 10,000 random iterations were performed for each scenario (subject match, gender match, HLA-risk match and microarray data points match) and differentially expressed genes that remained significant (p-values adjusted using false discovery rate) more than 90% of the time in all four scenarios were selected.

Genes	MA match	subject match	gender match	risk match
MAGI3	99.53%	97.56%	95.21%	98.31%
NEO1	99.2%	95.47%	93.27%	97.71%
SPRY1	100%	99.89%	99.86%	99.8%
PER3	100%	99.88%	99.84%	99.94%
CAMK2N1	99.65%	97.77%	96.66%	98.05%
SOX4	100%	99.98%	99.92%	99.98%
SELM	99.82%	98.33%	97.11%	98.72%
ADCY9	100%	99.87%	99.82%	99.92%
DCHS1	99.85%	99.66%	99.3%	99.75%
CYB561	99.34%	97.4%	97.89%	98.5%
TMCO4	98.49%	96.25%	94.99%	96.41%
IGF2BP3	100%	99.95%	99.96%	100%
NR3C2	100%	99.99%	100%	100%
APOL3	99.93%	98.98%	98.1%	98.52%
SOAT1	99.45%	95.91%	92.81%	98.11%
MID2	100%	99.78%	99.62%	99.95%
ZNF711	99.49%	96.98%	95.77%	96.84%
ENPP3	98.74%	93.33%	94.15%	93.49%

PGM2L1	95.89%	91.78%	91.73%	91.9%
NDNF	98.43%	92.27%	91.01%	94.16%
NEDD4	99.99%	99.85%	99.44%	99.82%
DACH1	99.99%	99.77%	99.45%	99.89%
ZMAT4	99.96%	98.7%	98.94%	99.89%
ADAMTS10	96.97%	91.73%	91.75%	91.13%
ZBTB32	99.76%	98.91%	98.33%	99.16%
CLDN1	100%	99.98%	99.94%	99.99%
HEY2	98.94%	96.34%	95.67%	98.05%
ADRB2	98.07%	93.92%	94.12%	93.56%
ZBTB38	99.29%	95.57%	92.08%	95.5%
PLEKHA5	99.94%	99.59%	99.23%	99.69%
ASIC1	99.63%	97.81%	96.4%	97.87%
CACNA2D2	99.28%	95.09%	92.9%	94.89%
SPG20	99.98%	99.72%	99.24%	99.81%
TLR3	98.13%	94.75%	93.27%	96.01%
RIC3	99.71%	96.05%	93.6%	96.6%
MEX3B	96.88%	93.43%	92.29%	93.74%
MCOLN2	99.91%	98.29%	96.42%	97.64%
KRT73	96.45%	91.2%	92.47%	95.72%
IL15RA	99.61%	97.32%	95.13%	96.96%
CACHD1	99.99%	99.44%	98.94%	99.82%
AFF2	99.99%	99.86%	99.53%	99.9%
ACVR1C	100%	99.92%	99.89%	99.96%
CLIC2	99.7%	97.12%	96.95%	98.19%
TMEM30B	99.98%	99.78%	99.59%	99.97%
FCHSD1	98.78%	94.79%	92.33%	96.12%
ZNF462	99.63%	96.62%	93.45%	97.74%
RNF122	96.77%	91.4%	93.95%	93.29%
PTCH1	99.89%	95.37%	92.3%	98.07%
DPPA4	99.79%	98.54%	96.6%	98.82%
ADAM23	99.99%	99.91%	99.76%	99.9%
TENM1	100%	99.5%	98.93%	99.65%
MICA	99.85%	97.51%	97.81%	95.93%
ZNF532	99.89%	98.69%	97.76%	99.74%
ZNF714	98.7%	97.6%	96.15%	99.58%
AGAP1	99.39%	97.4%	96.18%	96.74%
SYDE2	99.23%	93.7%	92.66%	91.18%
ITM2A	99.82%	98.14%	97.71%	99.47%

SEL1L3	97.57%	91.12%	90.52%	93.98%
TSHZ2	100%	99.87%	99.68%	99.87%
PARD3	99.23%	93.97%	91.53%	97.3%
SERPINB9P1	96.06%	93.57%	92.09%	93.19%
PDE7B	99.27%	95.2%	92.14%	96.51%
EOGT	99.68%	97.01%	94.28%	97.5%
SLC25A4	96.67%	90.94%	90.76%	94.19%
ADHFE1	97.8%	93.58%	93.21%	97.24%
TSPYL5	99%	93.98%	93.45%	96.56%
L3MBTL4	99.09%	95.89%	96.7%	94.16%

Table S2: Corrected p-values for time, grp, cohort and grp:time effects related to differentially expressed genes. Gene description is provided in the last column.

Genes	**Corrected p-values (False Discovery Rate)**				**Description**
	time	*grp*	*cohort*	*grp:time*	
MAGI3	1.00E-07	8.70E-01	1.00E+00	3.85E-03	Membrane-associated guanylate kinase, WW and PDZ domain-containing protein 3
NEO1	0.00E+00	8.77E-01	1.00E+00	5.79E-03	Neogenin
SPRY1	7.99E-05	1.82E-01	1.00E+00	1.85E-03	Protein sprouty homolog 1
PER3	0.00E+00	4.66E-01	1.00E+00	2.79E-05	Period circadian protein homolog 3
CAMK2N1	0.00E+00	8.75E-01	1.00E+00	2.36E-03	Calcium/calmodulin-dependent protein kinase II inhibitor 1
SOX4	0.00E+00	9.48E-01	1.00E+00	5.60E-06	Transcription factor SOX-4
SELM	0.00E+00	6.92E-01	1.00E+00	5.77E-04	Selenoprotein M
ADCY9	0.00E+00	4.31E-01	1.00E+00	2.59E-04	Adenylate cyclase type 9
DCHS1	3.80E-06	9.38E-01	1.00E+00	4.96E-04	Protocadherin-16
CYB561	1.00E-07	9.78E-01	1.00E+00	2.43E-03	Cytochrome b561
TMCO4	4.30E-06	4.70E-01	1.00E+00	2.20E-02	Transmembrane and coiled-coil domain-containing protein 4

IGF2BP3	0.00E+00	5.65E-01	1.00E+00	1.80E-06	Insulin-like growth factor 2 mRNA-binding protein 3
NR3C2	0.00E+00	4.24E-01	4.07E-01	1.00E-07	Mineralocorticoid receptor
APOL3	1.20E-06	8.28E-01	1.00E+00	7.05E-04	Apolipoprotein L3
SOAT1	2.64E-05	9.65E-01	1.00E+00	5.28E-03	Sterol O-acyltransferase 1
MID2	0.00E+00	7.79E-01	1.00E+00	1.98E-05	Probable E3 ubiquitin-protein ligase MID2
ZNF711	0.00E+00	7.72E-01	1.00E+00	5.08E-03	Zinc finger protein 711
ENPP3	7.70E-06	8.20E-01	1.00E+00	5.28E-03	Ectonucleotide pyrophosphatase/phosphodiesterase family member 3
PGM2L1	1.17E-03	5.07E-01	1.00E+00	3.30E-02	Glucose 1,6-bisphosphate synthase
NDNF	1.31E-04	8.15E-01	1.00E+00	2.32E-02	Protein NDNF
NEDD4	0.00E+00	8.20E-01	1.00E+00	1.22E-04	E3 ubiquitin-protein ligase NEDD4
DACH1	0.00E+00	9.33E-01	1.00E+00	6.82E-05	Dachshund homolog 1
ZMAT4	5.97E-02	8.98E-01	1.00E+00	2.38E-03	Zinc finger matrin-type protein 4
ADAMTS 10	1.66E-02	8.07E-01	1.00E+00	2.34E-02	A disintegrin and metalloproteinase with thrombospondin motifs 10
ZBTB32	0.00E+00	9.38E-01	1.00E+00	1.16E-03	Zinc finger and BTB domain-containing protein 32
CLDN1	0.00E+00	4.24E-01	1.00E+00	4.18E-05	Claudin-1
HEY2	9.24E-03	8.98E-01	1.00E+00	5.28E-03	Hairy/enhancer-of-split related with YRPW motif protein 2
ADRB2	8.76E-04	4.85E-01	1.00E+00	2.20E-02	Beta-2 adrenergic receptor
ZBTB38	0.00E+00	6.21E-01	1.00E+00	6.30E-03	Zinc finger and BTB domain-containing protein 38
PLEKHA5	1.67E-03	9.35E-01	1.00E+00	1.02E-03	Pleckstrin homology domain-containing family A member 5
ASIC1	4.00E-07	9.49E-01	1.00E+00	3.58E-03	Acid-sensing ion channel 1

CACNA2D2	0.00E+00	9.44E-01	1.00E+00	7.14E-03	Voltage-dependent calcium channel subunit alpha-2/delta-2
SPG20	0.00E+00	6.48E-01	1.00E+00	2.79E-05	Spartin
TLR3	0.00E+00	5.48E-01	1.00E+00	1.68E-02	Toll-like receptor 3
RIC3	0.00E+00	8.49E-01	1.00E+00	1.03E-03	Protein RIC-3
MEX3B	1.98E-02	7.63E-01	1.00E+00	2.20E-02	RNA-binding protein MEX3B
MCOLN2	0.00E+00	9.61E-01	1.00E+00	4.96E-04	Mucolipin-2
KRT73	1.33E-01	4.25E-01	1.00E+00	4.77E-02	Keratin, type II cytoskeletal 73
IL15RA	1.01E-05	5.97E-01	1.00E+00	3.85E-03	Interleukin-15 receptor subunit alpha
CACHD1	0.00E+00	8.66E-01	1.00E+00	3.05E-04	VWFA and cache domain-containing pr...
AFF2	0.00E+00	7.47E-01	1.00E+00	3.78E-05	AF4/FMR2 family member 2
ACVR1C	0.00E+00	4.31E-01	1.00E+00	4.42E-05	Activin receptor type-1C
CLIC2	1.81E-03	4.93E-01	1.00E+00	5.28E-03	Chloride intracellular channel protein 2
TMEM30B	0.00E+00	9.27E-01	1.00E+00	4.18E-05	Cell cycle control protein 50B
FCHSD1	0.00E+00	4.67E-01	1.00E+00	1.12E-02	F-BAR and double SH3 domains protein 1
ZNF462	1.00E-07	6.85E-01	1.00E+00	2.38E-03	Zinc finger protein 462
RNF122	6.76E-03	9.33E-01	1.00E+00	2.34E-02	RING finger protein 122
PTCH1	1.20E-05	1.36E-02	1.00E+00	1.87E-01	Protein patched homolog 1
DPPA4	1.00E-07	9.59E-01	1.00E+00	1.82E-03	Developmental pluripotency-associated protein 4
ADAM23	0.00E+00	9.89E-01	1.00E+00	2.79E-05	Disintegrin and metalloproteinase domain-containing protein 23
TENM1	0.00E+00	8.42E-01	1.00E+00	6.98E-05	Teneurin-1
MICA	1.83E-05	6.79E-01	1.00E+00	3.03E-03	MHC class I polypeptide-related sequence A
ZNF532	0.00E+00	8.84E-01	1.00E+00	1.02E-03	Zinc finger protein 532
ZNF714	4.04E-01	7.41E-01	1.00E+00	9.41E-03	Zinc finger protein 714
AGAP1	0.00E+00	6.75E-01	1.00E+00	5.75E-03	Arf-GAP with GTPase, ANK repeat and PH

					domain-containing protein 1
SYDE2	0.00E+00	8.12E-01	1.00E+00	3.08E-03	Rho GTPase-activating protein SYDE2
ITM2A	3.44E-05	8.94E-01	1.00E+00	2.97E-03	Amyloid beta A4 protein
SEL1L3	7.00E-07	4.96E-01	1.00E+00	3.30E-02	Protein sel-1 homolog 3
TSHZ2	0.00E+00	6.90E-01	1.00E+00	2.79E-05	Teashirt homolog 2
PARD3	2.00E-07	8.67E-01	1.00E+00	8.45E-03	Partitioning defective 3 homolog
SERPINB9P1	7.80E-02	8.60E-01	1.00E+00	3.76E-02	SERPINB9P1
PDE7B	1.20E-05	6.93E-01	1.00E+00	6.78E-03	cAMP-specific 3',5'-cyclic phosphodiesterase 7B
EOGT	0.00E+00	7.62E-01	1.00E+00	1.82E-03	EGF domain-specific O-linked N-acetylglucosamine transferase
SLC25A4	1.00E-07	6.13E-01	1.00E+00	2.34E-02	ADP/ATP translocase 1
ADHFE1	3.20E-06	8.38E-01	1.00E+00	2.34E-02	Hydroxyacid-oxoacid transhydrogenase, mitochondrial
TSPYL5	5.48E-04	8.23E-01	1.00E+00	6.24E-03	Ubiquitin carboxyl-terminal hydrolase 7
L3MBTL4	0.00E+00	8.94E-01	1.00E+00	4.38E-03	Lethal(3)malignant brain tumor-like protein 4

Table S3: Testing the effect of age as a covariate in seroconvertors and in non-seroconvertors

We fitted a linear model for each gene as a function of age in either seroconvertors or non seroconvertors. P-values adjusted for multiple correction using false discovery rate method are reported. The effect of age was declared significant in seroconvertors (ABP) for one single gene (NEO1, highlighted). Conversely, all genes had a significant age effect in non-seroconvertors (ABN) with the exception of ZMAT4, ADAMTS10, HEY2, MEX3B, KRT73, RNF122, ZNF714 and SERPINB9P1 (highlighted).

Genes	Adjusted p-values (p_{adj})

	Age effect in ABP	Age effect in ABN
MAGI3	6.76E-01	9.97E-06
NEO1	4.49E-02	9.93E-13
SPRY1	1.00E+00	2.11E-03
PER3	9.53E-01	1.16E-10
CAMK2N1	1.79E-01	2.50E-12
SOX4	4.22E-01	5.79E-14
SELM	5.02E-01	4.08E-14
ADCY9	9.67E-01	1.15E-11
DCHS1	9.29E-01	3.14E-04
CYB561	9.86E-01	2.65E-06
TMCO4	1.00E+00	4.10E-04
IGF2BP3	1.49E-01	1.24E-17
NR3C2	6.42E-01	8.92E-21
APOL3	1.00E+00	9.14E-05
SOAT1	9.96E-01	1.79E-03
MID2	9.05E-02	4.02E-22
ZNF711	8.44E-01	3.45E-09
ENPP3	9.94E-01	1.64E-04
PGM2L1	1.00E+00	2.11E-02
NDNF	1.00E+00	3.57E-03
NEDD4	9.04E-01	4.42E-06
DACH1	9.05E-02	8.56E-13
ZMAT4	1.00E+00	9.58E-02
ADAMTS10	1.00E+00	5.98E-02
ZBTB32	9.98E-01	1.02E-08
CLDN1	1.00E+00	1.44E-07
HEY2	1.00E+00	9.02E-02
ADRB2	1.00E+00	6.98E-03
ZBTB38	8.67E-01	1.02E-08
PLEKHA5	1.00E+00	1.67E-02
ASIC1	7.78E-01	1.27E-05
CACNA2D2	4.33E-01	3.08E-08
SPG20	9.40E-01	3.46E-07
TLR3	8.78E-01	4.19E-07
RIC3	7.70E-01	7.45E-12
MEX3B	1.00E+00	7.67E-02
MCOLN2	7.99E-01	4.76E-10
KRT73	1.00E+00	3.93E-01

IL15RA	1.00E+00	2.77E-04
CACHD1	1.67E-01	2.05E-07
AFF2	3.45E-01	4.24E-09
ACVR1C	9.02E-01	1.10E-11
CLIC2	1.00E+00	1.95E-03
TMEM30B	3.97E-01	6.22E-12
FCHSD1	9.65E-01	2.36E-07
ZNF462	7.30E-01	1.21E-05
RNF122	1.00E+00	7.33E-02
PTCH1	1.00E+00	5.60E-04
DPPA4	9.64E-01	2.53E-05
ADAM23	1.79E-01	1.50E-11
TENM1	1.49E-01	8.56E-13
MICA	1.00E+00	4.13E-04
ZNF532	9.64E-01	2.21E-10
ZNF714	1.00E+00	7.28E-01
AGAP1	7.31E-01	3.71E-07
SYDE2	1.49E-01	8.56E-13
ITM2A	9.19E-01	8.86E-04
SEL1L3	1.00E+00	1.24E-04
TSHZ2	3.19E-01	7.65E-17
PARD3	6.89E-01	1.84E-05
SERPINB9P1	9.98E-01	2.16E-01
PDE7B	4.87E-01	3.14E-04
EOGT	2.83E-01	1.44E-07
SLC25A4	9.96E-01	5.13E-05
ADHFE1	9.98E-01	4.10E-04
TSPYL5	8.86E-01	1.95E-03
L3MBTL4	6.76E-01	5.21E-09

Table S4. Gene ontology (GO) terms associated with protein interaction network uncovered ubiquitin related GO processes as significantly enriched.

GOID	**GOTerm**	**# Genes**	**% Associated Genes**	**Term PValue**	**ste**
GO:0006511	ubiquitin-dependent protein catabolic process	24	10.256411	8.67035E-12	3.99E-09

GO:0019941	modification-dependent protein catabolic process	24	10.126582	1.14081E-11	5.24E-09
GO:0043632	modification-dependent macromolecule catabolic process	24	10	1.49465E-11	6.85E-09
GO:0051603	proteolysis involved in cellular protein catabolic process	26	8.99654	2.16606E-11	9.90E-09
GO:0030163	protein catabolic process	31	7.416268	3.11478E-11	1.42E-08
GO:0010558	negative regulation of macromolecule biosynthetic process	43	5.541237	4.01054E-11	1.82E-08
GO:0044257	cellular protein catabolic process	27	8.385093	4.37347E-11	1.99E-08
GO:0009890	negative regulation of biosynthetic process	44	5.392157	5.47481E-11	2.48E-08
GO:0010605	negative regulation of macromolecule metabolic process	57	4.4392524	7.66896E-11	3.47E-08
GO:0051246	regulation of protein metabolic process	60	4.2613635	1.03037E-10	4.65E-08
GO:0031327	negative regulation of cellular biosynthetic process	43	5.354919	1.22267E-10	5.50E-08
GO:2000113	negative regulation of cellular macromolecule biosynthetic process	41	5.5181694	1.50039E-10	6.74E-08
GO:0051172	negative regulation of nitrogen compound metabolic process	44	5.1825676	2.02456E-10	9.07E-08
GO:0016567	protein ubiquitination	27	7.803468	2.26658E-10	1.01E-07
GO:0023051	regulation of signaling	60	4.1899443	2.50705E-10	1.12E-07
GO:0032446	protein modification by small protein conjugation	28	7.3878627	3.57998E-10	1.59E-07
GO:0010646	regulation of cell communication	59	4.1696115	4.52709E-10	2.01E-07

GO:0031324	negative regulation of cellular metabolic process	56	4.242424	7.43378E-10	3.29E-07
GO:0009966	regulation of signal transduction	57	4.215976	8.7006E-10	3.85E-07
GO:0051247	positive regulation of protein metabolic process	44	4.894327	1.2826E-09	5.66E-07
GO:0032268	regulation of cellular protein metabolic process	55	4.230769	2.07519E-09	9.13E-07
GO:0009892	negative regulation of metabolic process	58	4.008293	3.20088E-09	1.41E-06
GO:0031399	regulation of protein modification process	44	4.7413793	3.49392E-09	1.53E-06
GO:0035556	intracellular signal transduction	51	4.336735	3.60638E-09	1.58E-06
GO:0007167	enzyme linked receptor protein signaling pathway	26	7.0652175	4.31931E-09	1.88E-06
GO:0043161	proteasome-mediated ubiquitin-dependent protein catabolic process	17	10.493827	8.41862E-09	3.66E-06
GO:1902679	negative regulation of RNA biosynthetic process	34	5.4574637	1.08986E-08	4.73E-06
GO:0032270	positive regulation of cellular protein metabolic process	40	4.842615	1.2555E-08	5.44E-06
GO:0070647	protein modification by small protein conjugation or removal	29	6.1181436	1.31926E-08	5.70E-06
GO:0010498	proteasomal protein catabolic process	18	9.473684	1.52536E-08	6.57E-06
GO:0009057	macromolecule catabolic process	31	5.7513914	1.6446E-08	7.07E-06

GO:0007179	transforming growth factor beta receptor signaling pathway	13	13.978495	1.6602E-08	7.12E-06
GO:0007178	transmembrane receptor protein serine/threonine kinase signaling pathway	16	10.7382555	1.70506E-08	7.30E-06
GO:0044265	cellular macromolecule catabolic process	27	6.3679247	1.95744E-08	8.36E-06
GO:0045892	negative regulation of transcription, DNA-templated	33	5.4455447	1.98767E-08	8.47E-06
GO:1903507	negative regulation of nucleic acid-templated transcription	33	5.4365735	2.06965E-08	8.80E-06
GO:0071495	cellular response to endogenous stimulus	30	5.791506	2.50185E-08	1.06E-05
GO:0031647	regulation of protein stability	18	9.090909	2.91544E-08	1.23E-05
GO:0071560	cellular response to transforming growth factor beta stimulus	14	12.068966	3.1442E-08	1.33E-05
GO:0071310	cellular response to organic substance	42	4.5801525	3.36125E-08	1.42E-05
GO:0045934	negative regulation of nucleobase-containing compound metabolic process	37	4.9071617	3.66044E-08	1.54E-05
GO:0071559	response to transforming growth factor beta	14	11.864407	3.92281E-08	1.64E-05
GO:0006508	proteolysis	37	4.8748355	4.35626E-08	1.82E-05
GO:0051253	negative regulation of RNA metabolic process	34	5.1282053	5.15846E-08	2.15E-05
GO:0010033	response to organic substance	48	4.155844	5.73892E-08	2.39E-05
GO:0010629	negative regulation of gene expression	40	4.602992	7.65575E-08	3.18E-05

GO:0007166	cell surface receptor signaling pathway	42	4.402516	8.20833E-08	3.40E-05
GO:0023057	negative regulation of signaling	32	5.2287583	8.88389E-08	3.67E-05
GO:0009968	negative regulation of signal transduction	31	5.254237	1.32137E-07	5.44E-05
GO:0006913	nucleocytoplasmic transport	18	8.144796	1.58243E-07	6.50E-05
GO:0051169	nuclear transport	18	8.035714	1.94045E-07	7.96E-05
GO:0048585	negative regulation of response to stimulus	35	4.7233467	2.41517E-07	9.88E-05
GO:0010648	negative regulation of cell communication	31	5.107084	2.49885E-07	1.02E-04
GO:0006468	protein phosphorylation	41	4.2399173	3.96823E-07	1.62E-04
GO:1901575	organic substance catabolic process	36	4.494382	6.33467E-07	2.57E-04
GO:1903320	regulation of protein modification by small protein conjugation or removal	14	9.459459	6.79392E-07	2.75E-04
GO:0031396	regulation of protein ubiquitination	13	9.923664	1.00167E-06	4.05E-04
GO:0032269	negative regulation of cellular protein metabolic process	27	5.222437	1.10798E-06	4.47E-04
GO:0034613	cellular protein localization	38	4.284104	1.12469E-06	4.52E-04
GO:0070727	cellular macromolecule localization	38	4.269663	1.17515E-06	4.71E-04
GO:0043067	regulation of programmed cell death	33	4.5643153	1.7922E-06	7.17E-04
GO:0016458	gene silencing	10	12.820513	1.80914E-06	7.22E-04
GO:0017015	regulation of transforming growth factor beta receptor signaling pathway	9	14.754098	1.81984E-06	7.24E-04

GO:0000122	negative regulation of transcription from RNA polymerase II promoter	22	5.866667	1.89199E-06	7.51E-04
GO:0045862	positive regulation of proteolysis	15	8.1081085	1.93071E-06	7.65E-04
GO:1903844	regulation of cellular response to transforming growth factor beta stimulus	9	14.5161295	2.09487E-06	8.27E-04
GO:0044248	cellular catabolic process	37	4.176072	2.48448E-06	9.79E-04
GO:0007183	SMAD protein complex assembly	5	38.46154	2.56968E-06	1.01E-03
GO:0008219	cell death	39	4.09234	2.68192E-06	1.05E-03
GO:0043069	negative regulation of programmed cell death	23	5.5555553	2.73923E-06	1.07E-03
GO:0006898	receptor-mediated endocytosis	11	10.891089	2.822E-06	1.10E-03
GO:0010941	regulation of cell death	34	4.3202033	3.06065E-06	1.19E-03
GO:0042981	regulation of apoptotic process	32	4.4880786	3.38371E-06	1.31E-03
GO:0031400	negative regulation of protein modification process	19	6.291391	3.68847E-06	1.43E-03
GO:0012501	programmed cell death	36	4.157044	3.91636E-06	1.51E-03
GO:0051649	establishment of localization in cell	35	4.1666665	5.45021E-06	2.09E-03
GO:0051248	negative regulation of protein metabolic process	27	4.8042703	5.4476E-06	2.10E-03
GO:1903322	positive regulation of protein modification by small protein conjugation or removal	10	11.235955	6.11995E-06	2.34E-03
GO:0006915	apoptotic process	35	4.1273584	6.41573E-06	2.45E-03
GO:0046907	intracellular transport	31	4.403409	6.78349E-06	2.58E-03

GO:0043066	negative regulation of apoptotic process	22	5.4187193	6.91916E-06	2.63E-03
GO:0030162	regulation of proteolysis	20	5.681818	9.39445E-06	3.56E-03
GO:0060968	regulation of gene silencing	7	17.073172	9.83062E-06	3.72E-03
GO:0031398	positive regulation of protein ubiquitination	9	11.842105	1.17441E-05	4.43E-03
GO:0060548	negative regulation of cell death	23	5.077263	1.22433E-05	4.60E-03
GO:0071363	cellular response to growth factor stimulus	17	6.115108	1.80353E-05	6.76E-03
GO:0044403	symbiosis, encompassing mutualism through parasitism	20	5.4200544	1.88003E-05	7.03E-03
GO:0043112	receptor metabolic process	10	9.615385	2.46538E-05	9.20E-03
GO:0070848	response to growth factor	17	5.923345	2.72112E-05	1.01E-02
GO:0090092	regulation of transmembrane receptor protein serine/threonine kinase signaling pathway	10	9.433962	2.91313E-05	1.08E-02
GO:0001932	regulation of protein phosphorylation	30	4.178273	3.38955E-05	1.25E-02
GO:0098662	inorganic cation transmembrane transport	14	6.6037736	4.46889E-05	1.65E-02
GO:0060765	regulation of androgen receptor signaling pathway	5	22.727272	4.58175E-05	1.69E-02
GO:0009755	hormone-mediated signaling pathway	9	9.89011	5.07691E-05	1.86E-02
GO:0031401	positive regulation of protein modification process	27	4.245283	5.39201E-05	1.97E-02
GO:0014070	response to organic cyclic compound	17	5.610561	5.41459E-05	1.97E-02

GO:1902895	positive regulation of pri-miRNA transcription from RNA polymerase II promoter	4	33.333332	5.39081E-05	1.97E-02
GO:0097190	apoptotic signaling pathway	18	5.3254437	6.38891E-05	2.32E-02
GO:0031958	corticosteroid receptor signaling pathway	3	60	6.43326E-05	2.33E-02
GO:0038092	nodal signaling pathway	3	60	6.43326E-05	2.33E-02
GO:0042921	glucocorticoid receptor signaling pathway	3	60	6.43326E-05	2.33E-02
GO:0007399	nervous system development	20	4.9261084	7.32176E-05	2.64E-02
GO:0044092	negative regulation of molecular function	25	4.363002	7.63324E-05	2.75E-02
GO:0040029	regulation of gene expression, epigenetic	9	9.278351	8.40722E-05	3.02E-02
GO:0016032	viral process	18	5.142857	0.000100086	3.58E-02
GO:0030518	intracellular steroid hormone receptor signaling pathway	8	10.126582	0.000113406	4.04E-02
GO:0098655	cation transmembrane transport	14	6.060606	0.000113198	4.04E-02
GO:0098660	inorganic ion transmembrane transport	14	6.060606	0.000113198	4.04E-02
GO:0007346	regulation of mitotic cell cycle	15	5.7471266	0.000116414	4.13E-02
GO:0050821	protein stabilization	10	8	0.000119386	4.23E-02
GO:2001233	regulation of apoptotic signaling pathway	14	5.982906	0.000129833	4.58E-02
GO:0042787	protein ubiquitination involved in ubiquitin-dependent protein catabolic process	7	11.4754095	0.000139067	4.90E-02
GO:0006623	protein targeting to vacuole	4	26.666666	0.000142214	4.99E-02

GO:0072666	establishment of protein localization to vacuole	4	26.666666	0.000142214	4.99E-02

Table S5. Genes enriched in the GSEA pathways shown in Table 1.

GSEA pathway	**Genes Involved**
T-cell activation	MAGI3, NEO1, SPRY1, CAMK2N1, SOX4, SELM, DCHS1, CYB561, IGF2BP3, NR3C2, APOL3, ZNF711, CLDN1, HEY2, ADRB2, TLR3, MEX3B, IL15RA, CACHD1, AFF2, ACVR1C, CLIC2, RNF122, PTCH1, DPPA4, ADAM23, MICA, AGAP1, SYDE2, ITM2A, SLC25A4, TSPYL5
Naive CD8 T cell genes	SOX4, DCHS1, IGF2BP3, NR3C2, ADRB2, PTCH1, ADAM23, ITM2A, SEL1L3, TSPYL5
Thymocyte development	NEO1, SPRY1, PER3, CAMK2N1, SOX4, SELM, ADCY9, CYB561, IGF2BP3, NR3C2, SOAT1, MID2, ZNF711, PGM2L1, NEDD4, DACH1, ZMAT4, ZBTB32, CLDN1, ADRB2, ZBTB38, PLEKHA5, CACNA2D2, SPG20, TLR3, MCOLN2, KRT73, IL15RA, ACVR1C, TMEM30B, ZNF462, PTCH1, ADAM23, MICA, ZNF532, AGAP1, SYDE2, ITM2A, SEL1L3, PDE7B, SLC25A4, ADHFE1, TSPYL5
Toll-like receptor signalling in DCs	MAGI3, SOX4, ADCY9, IGF2BP3, ZNF711, ENPP3, PGM2L1, DACH1, ZMAT4, ADAMTS10, ADRB2, MEX3B, MCOLN2, CACHD1, CLIC2, FCHSD1, PTCH1,

	DPPA4, ZNF532, AGAP1, ITM2A, PDE7B, SLC25A4, TSPYL5
Pathogen response	SOX4, ADCY9, SOAT1, NEDD4, ITM2A
Genes down regulated in Foxp3+ T cells	MAGI3, NEO1, SPRY1, PER3, CAMK2N1, SOX4, SELM, ADCY9, DCHS1, CYB561, IGF2BP3, APOL3, MID2, ZNF711, PGM2L1, NEDD4, ZMAT4, PLEKHA5, CACNA2D2, TLR3, MEX3B, MCOLN2, IL15RA, PTCH1, DPPA4, ADAM23, ZNF532, AGAP1, ITM2A, SLC25A4, TSPYL5
Macrophage differentiation	CAMK2N1, SELM, SPG20, ACVR1C, SYDE2
Memory CD4 T cell genes	MAGI3, SPRY1, SOX4, ADCY9, TMCO4, NR3C2, SOAT1, ENPP3, NEDD4, ZBTB32, ADRB2, CACNA2D2, RIC3, IL15RA, CACHD1, ACVR1C, DPPA4, ZNF532, ITM2A, SEL1L3, PDE7B, SLC25A4, TSPYL5
Naive B cell genes	MAGI3, CAMK2N1, SOX4, DCHS1, CYB561, TMCO4, NR3C2, APOL3, MID2, NEDD4, ZBTB32, CLDN1, ADRB2, ZBTB38, PLEKHA5, CACNA2D2, SPG20, TLR3, MCOLN2, IL15RA, TMEM30B, PTCH1, AGAP1, ITM2A, SEL1L3, PDE7B
Type I IFN response	MAGI3, SOX4, APOL3, ZNF711, DACH1, PLEKHA5, TLR3, IL15RA

B cell lineage	NEO1, SPRY1, DCHS1, CYB561, IGF2BP3, NR3C2, MID2, ZNF711, DACH1, ZMAT4, ADRB2, ZBTB38, TLR3, MCOLN2, RNF122, PTCH1, MICA, ZNF532, ITM2A, SEL1L3, TSHZ2, PDE7B, TSPYL5
Naive CD4 T cell genes	CAMK2N1, ADCY9, IGF2BP3, NEDD4, ZBTB32, ZBTB38, PLEKHA5, CACNA2D2, RIC3, CLIC2, PTCH1, ZNF532, ITM2A, SLC25A4
Memory CD8 T cell genes	MAGI3, SPRY1, SOX4, ADCY9, IGF2BP3, MID2, PGM2L1, NEDD4, DACH1, ADAMTS10, ZBTB32, CLDN1, ADRB2, PLEKHA5, TLR3, MEX3B, MCOLN2, IL15RA, CACHD1, PTCH1, ADAM23, SYDE2, SEL1L3, SLC25A4, ADHFE1, TSPYL5
Viral response	ZBTB38, TLR3, IL15RA, ZNF532
Toll-like receptor signalling in macrophages	DCHS1, RIC3, ZNF532, TSHZ2
CD4 T cell lineage	CYB561, PGM2L1, PLEKHA5, MCOLN2
Interferon gamma response	SOX4, NR3C2, APOL3, ZNF711, ZBTB38, PLEKHA5, MCOLN2, KRT73, IL15RA, TMEM30B, MICA, ITM2A, SEL1L3
IL-4 induced genes	SOX4, APOL3, ZNF711, RNF122

B-cell receptor signalling	SPRY1, PER3, CAMK2N1, CYB561, ZMAT4, IL15RA, TMEM30B
Macrophage activation	MAGI3, SPRY1, SOX4, CYB561, NR3C2, ADRB2, PLEKHA5, TLR3, RIC3, RNF122, DPPA4, ZNF532, ITM2A, TSPYL5
Memory B cell genes	NEO1, NR3C2, ZBTB32, MCOLN2, TMEM30B, PDE7B, TSPYL5
CD8 T cell lineage	MAGI3, CAMK2N1, CLDN1, PLEKHA5, TLR3, MCOLN2, IL15RA, PTCH1, ADAM23, ITM2A
Pathogen response in DCs	SOX4, ADCY9, DCHS1, NEDD4, PLEKHA5, ACVR1C, PTCH1, SYDE2, PARD3
Epithelial cells	PER3, SOX4, CYB561, DACH1, DPPA4, SEL1L3, PDE7B
Germinal centre B cells	CYB561, IGF2BP3, DACH1, CACNA2D2, ACVR1C, CLIC2, TMEM30B, ADAM23, ITM2A
B-cell development	CAMK2N1, SOX4, TMCO4, APOL3, ENPP3, PGM2L1, NEDD4, HEY2, PLEKHA5, CACNA2D2, TLR3, MEX3B, MCOLN2, IL15RA, AGAP1
CD8+ DC lineage	MCOLN2, IL15RA, AFF2, TSPYL5
Hematopoietic stem cell regulation	IGF2BP3, ADAMTS10, TLR3, MCOLN2, TMEM30B, DPPA4, MICA, PARD3

Monocyte lineage genes	NEO1, SPRY1, CAMK2N1, SOX4, SELM, ADCY9, CYB561, TMCO4, IGF2BP3, NR3C2, APOL3, SOAT1, MID2, PGM2L1, NEDD4, ZBTB32, ADRB2, PLEKHA5, SPG20, TLR3, MCOLN2, IL15RA, ACVR1C, CLIC2, SYDE2, ITM2A, PARD3, SLC25A4
Macrophage development	MAGI3, SPRY1, CAMK2N1, SOX4, ADCY9, IGF2BP3, ZNF711, ENPP3, NEDD4, DACH1, CLDN1, ZBTB38, PLEKHA5, CACNA2D2, SPG20, KRT73, CACHD1, AFF2, ZNF462, PTCH1, ZNF532, SYDE2, ITM2A, PDE7B, L3MBTL4
DC activation	PER3, SELM, APOL3, SOAT1, MCOLN2, IL15RA, AGAP1
PBMC expressed genes	CLDN1, TLR3, IL15RA, FCHSD1
T-cell development	MAGI3, SOX4, SELM, IGF2BP3, MID2, TLR3, IL15RA, PTCH1, MICA, ZNF532, SYDE2, ITM2A, SEL1L3, PDE7B, ADHFE1
Homeostatically converted Treg cells	SPRY1, IGF2BP3, NR3C2, NEDD4, ZBTB32, ZBTB38, ITM2A, TSPYL5
Myeloid-derived suppressor cells	PER3, DCHS1, CYB561, MID2, ZMAT4, TLR3, IL15RA, TMEM30B, ZNF462, ADAM23, ITM2A, PDE7B
Plasmacytoid DC lineage	CLDN1, ADRB2, TLR3

B1 B cell genes	IL15RA, AGAP1, ITM2A, SEL1L3, PARD3
Th1 T cell lineage	NEO1, SPRY1, CAMK2N1, MID2, ADRB2, CACNA2D2, TLR3, MCOLN2, PTCH1, SYDE2, SLC25A4
NK cell maturation	ADRB2, FCHSD1, ZNF532
Follicular helper CD4 T cells	MAGI3, PER3, CAMK2N1, ADCY9, NR3C2, DACH1
Mast cell activation	CYB561, ENPP3, ZNF462
Plasma cell genes	MAGI3, CACNA2D2, SYDE2
Treg cells	APOL3, ENPP3, ADRB2, ZBTB38, TLR3, RIC3, CLIC2, ADAM23, ZNF532, ITM2A, SLC25A4
Pro-B cell genes	CYB561, TLR3, SEL1L3
Genes down regulated after macrophage activation	SOX4, ADCY9, ADAMTS10, MEX3B, MICA, ITM2A
Retinoic acid signalling in DCs	NEO1, NR3C2, ADAMTS10, PLEKHA5, TLR3, IL15RA
Genes down-regulated by IL-4	PGM2L1, ADRB2, ZNF462
Bone marrow progenitors	SELM, CYB561, FCHSD1

Reactive oxygen species induced genes in DCs	PER3, ZBTB38, MCOLN2
Effector memory T cell lineage	PER3, DCHS1, MID2, PGM2L1, ZBTB38, MCOLN2, CLIC2, PTCH1, DPPA4
Peripheral blood B cell genes	SPRY1, APOL3, ZBTB32, ZBTB38, MEX3B, IL15RA, TMEM30B, ZNF532, SLC25A4
Genes down-regulated in Treg	ZBTB32, ZBTB38, SLC25A4
IL-21 response genes	ZBTB38, RIC3, ADHFE1
Lung DCs	MAGI3, IL15RA, AFF2, PTCH1, ZNF532, SLC25A4
Plasmablast genes	PER3, NEDD4, TLR3
Aryl-hydrocarbon receptor signalling in DCs	NEO1, SPRY1, PER3, SOAT1, NEDD4, ZBTB32, MCOLN2, RNF122, PTCH1, ADAM23, SLC25A4
CD8 negative DCs	ADRB2, ITM2A, SLC25A4
Genes down-regulated by retinoic acid in CD4 T cells	NEO1, SPRY1, SELM, IGF2BP3, NR3C2, NEDD4, ZBTB32, CLDN1, ADRB2, IL15RA, PTCH1, MICA, ITM2A, SEL1L3, SLC25A4, L3MBTL4
Monocyte activation	IGF2BP3, NEDD4, ITM2A
Neuronal cell genes	NR3C2, ADRB2, IL15RA, CACHD1, AFF2, ACVR1C, CLIC2, MICA, SYDE2

Induced Treg genes	DCHS1, PTCH1, ADAM23
T follicular helper cell genes	MCOLN2, CACHD1, ADHFE1
Tumour monocytes	NR3C2, ZNF711, ZNF532
Peritoneal cavity B cells	MAGI3, SOX4, MCOLN2
Toll-like receptor signalling in B-cells	SOX4, TMCO4, ITM2A
CD8 T cell tumour response	SOX4, TMCO4, NEDD4, ADRB2, ACVR1C, ITM2A
Tissue resident DCs	PER3, HEY2, TLR3
Pathogen response in macrophages	CACNA2D2, SPG20, PTCH1
Eosinophil response	PGM2L1, ACVR1C, SLC25A4
Neutrophil response	SOX4, MCOLN2, SEL1L3
Conventional dendritic cells	PTCH1, ZNF532, AGAP1
Hematopoietic stem cells	CAMK2N1, IGF2BP3, MID2, TMEM30B, SYDE2
Genes up-regulated in Foxp3+ T cells	CAMK2N1, NR3C2, NEDD4, CACNA2D2, ITM2A, SLC25A4
Th17 cell differentiation	CAMK2N1, SOX4, SELM
Mast cell genes	CAMK2N1, CYB561, NEDD4

Monocyte derived DC genes	SOX4, ADCY9, SOAT1, DACH1, ZMAT4, ADAMTS10, CLDN1, ADRB2, ZBTB38, TLR3, MEX3B, KRT73, FCHSD1, ZNF462, PDE7B
NF-kB repressed genes in DCs	MAGI3, TLR3, IL15RA, PTCH1
Fat-tissue T cells	SOX4, IGF2BP3, ZNF711, PGM2L1, DACH1, ZMAT4, PDE7B
Viral RNA response	PER3, NR3C2, MID2
IL-6 response genes	ENPP3, NEDD4, DACH1, CLDN1, ZNF462, L3MBTL4
Genes down-regulated by TGF-beta	CAMK2N1, ADCY9, ZNF711, PLEKHA5, SPG20, KRT73, CACHD1, AFF2, ZNF532