SUPPLEMENTAL METHODS

PSEA analysis

Gene expression data set

The dataset included whole tissue transcriptomes from a total of 310 gingival tissue samples obtained from 120 patients with periodontitis, comprising interproximal papillae that were either periodontitis-affected [241 samples showing bleeding on probing (BoP), probing depth (PD) ≥4mm and clinical attachment level (CAL≥3mm)] or clinically healthy (69 samples; with no BoP, PD≤4 mm, and CAL≤2 mm). Inclusion criteria, demographics, sample characteristics and processing pipeline have been published previously (Kebschull et al. 2013).

Five outlier arrays (all from periodontitis-affected samples) were identified by the GNUSE method and removed, as earlier described (McCall et al. 2014). Data was first normalized using GCRMA (Wu et al. 2004; Wu and Irizarry 2005) and further between batches using COMBAT (Johnson et al. 2007) implemented in the SVA R package, as earlier described (Sawle et al. 2016). Whenever two periodontitis-affected samples were available from the same donor, their intensities were averaged. Thus, the dataset further analyzed comprised 118 periodontitis-associated and 69 healthy gingival tissue samples.

Summary of the Barcode Methodology

The Gene Expression Barcode 3.0 method (McCall et al. 2011; McCall et al. 2014) takes as its input the expression of each probeset for a given chip platform from samples from a wide variety of cell and tissue types, normalized by the frozen Robust Multichip Algorithm

(fRMA; McCall et al. 2010). The method infers an expression threshold for each chip by an hierarchical (Gelman and Hill 2006) Bayesian (Gelman et al. 2014) mixture model (Everitt and Hand 1981) which is an adaptation of the Probability of Expression model (POE; Parmigiani et al. 2002). The probability of expression of a given probeset in each cell or tissue type is then taken to be the fraction of samples of that cell or tissue type whose expression is greater than or equal to the expression threshold of that probeset.

Filtering of marker probesets (genes)

The list of marker probesets were filtered further based upon gene expression data as follows: (i) if two probesets representing the same gene had a Pearson correlation coefficient of<0.7, one of them was discarded; (ii) if two probesets which were candidate markers for different cell types had a Pearson correlation coefficient p-value<0.05, one was eliminated (this criterion guaranteed that the markers for different cell types were not correlated in our models); (iii) the variance inflation factor (VIF) described below for all of the probesets within a marker for a cell type was<10; (iv) probesets were further discarded based upon known expression in other cellular subtypes in the gingiva.

Initial filtering of probesets in the differential expression analysis

Probesets were removed from the differential expression analyses based upon the following criteria listed in the PSEA publication (Kuhn et al. 2011) : (i) probesets used as markers, to avoid circular reasoning; (ii) probesets with an adjusted coefficient of determination

R^2<0.6 for the best model, indicating a poor fit; and (iii) probesets whose intercepts were>0.5 their average expression, indicating that their expression did not vary with the marker for a cell type.

Cell type-specific filtering of probesets in the differential expression analysis

Additional filtering criteria were applied to select probesets of interest: (i) p-value of presence of the gene in the cell type that is differentially expressed<0.05; (ii) p-value of differential expression in periodontitis-affected versus healthy gingiva<0.05; (iii) absolute log2FC differential expression of>0.4, and (iv) confidence coefficient (CC); i.e., the fraction of models with the same cell-type differentially expressed within 2 AIC of the best model, =1. Additional filtered-out probesets included those with negative coefficients of presence (corresponding to negative concentration) and those whose differential expression led to net negative concentration. Redundant probesets for the same gene were removed.

Validation of PSEA predicted genes

Gingival tissue harvesting and preparation of single cell suspensions

Gingival tissue samples were harvested from patients in conjunction with periodontal surgical procedures (pocket elimination/reduction surgery, crown lengthening or tooth extraction) after approval by the Columbia University Medical Center Institutional Review Board (Protocol # AAAR0526). Patients were recruited at the Clinic of Graduate Periodontics of the College of Dental Medicine and informed consent was obtained. All patients were systemically healthy, non-pregnant, non-smokers who had not used antibiotics or anti-

inflammatory drugs for the preceding 3-month period, as in our previous publications (Kebschull et al. 2013). Interproximal papillae included in the surgical area were harvested and originated either from areas affected by periodontitis, i.e., an interproximal site with PD ≥ 5 mm, with concomitant CAL ≥ 3 mm, presence of radiographic bone, and BoP or from clinically healthy sites, (i.e., sites with PD≤ 3 mm, no CAL, no radiographic bone loss and no BoP). From each patient one periodontitis-affected gingival tissue sample and one healthy tissue sample were harvested (15 pairs, n=30). Gingival tissue samples were processed to form a single-cell suspension using an established laboratory protocol (Almubarak et al. 2020), using a commercially available tissue dissociation kit (Miltenyi Biotech, USA). Samples were kept in Dulbecco's Phosphate-Buffered Saline, and minced into small pieces after washing with saline. Samples were processed in c-tubes (Miltenyi Biotech, USA) which contained 2.35 ml of RPMI 1640 (Thermo Fisher, USA) and 100 µl of Enzyme H, 50 µl of Enzyme R, and 12.5 µl of Enzyme A (Miltenyi Biotech, USA), as recommended by the manufacturer. A gentle MACS Dissociator (Miltenyi Biotech, USA) was used for tissue disruption and enzymatic digestion at 37°C as recommended by the manufacturer. The combination of this mechanical and enzymatic digestion leads to formation single cell suspensions, while maintaining cellular integrity. Cell suspensions were filtered with 70-micrometer filters and each sample was washed with 15 ml of RPMI 1640 solution. Cells were pelleted by centrifugation for 7 minutes (relative centrifugal field: 0.4) and cryopreserved immediately. The pellet was mixed with a freezing medium, which contained 90% of fetal bovine serum (FBS; Corning, USA) and 10% of dimethyl sulfoxide (DMSO; Sigma-Aldrich, USA). The samples were transferred to isopropanol chambers in a -80°C freezer and then transferred into liquid nitrogen within 24h. This protocol has been shown to have minimal effects on the transcriptional profiles after cell revival (Guillaumet-Adkins et al. 2017).

Immuno-magnetic separation of epithelial cells and B cells

All cryopreserved samples were revived from liquid nitrogen and viability was assessed using Trypan blue staining using a TC20 Automated Cell Counter. Approximately 5-6 x 10^8 cells were aliquoted in 300 µl of autoMACS® Running Buffer (Miltenyi Biotec, USA) and 100 µl of FcR blocking reagent (Miltenyi Biotec, USA) was added and shaken gently for 5 minutes. 100 µl of CD326 (EpCAM) magnetic MicroBeads for isolation of epithelial cells or CD19 magnetic MicroBeads (Miltenyi Biotec, USA) for isolation of B cells were added to the suspension and incubated on a rocking platform for 30 minutes at 4°C. EpCAM is an established marker for epithelial cell isolation and has been reported to be highly expressed in gingival epithelial cells (Hasegawa et al. 2017; Huang et al. 2018; Balfe et al. 2018; Hyun et al. 2019). The cells were washed by adding 5 ml of buffer and centrifuged at 300×g for 10 minutes. The supernatants were aspirated and the cells were suspended in 500 µl of buffer. An LS column (Miltenyi Biotec, USA) was placed in the magnetic field and prepared by rinsing with 500µl of buffer. The cell suspension was applied gently from the sides onto the column and was washed three times using 500 µl of buffer. The column was removed from the separator and placed on a collection tube provided by the kit. 1ml of buffer was pipetted onto the column and the magnetically labeled cells were pushed by the plunger and isolated in a new tube. The isolated cells were pelleted by centrifugation at 800xg for 7 minutes, 500 µl TRI reagent (Zymo Research, USA) were added to the cell pellet and mixed well by pipetting, and the samples were kept at -80°C for RNA isolation.

RNA isolation

The Direct-zol RNA Microprep Kit (Zymo Research, USA) was used for RNA isolation according to the manufacturer's instructions. RNA was eluted in 30 µl RNase-free water. The quantity and quality of the RNA (260/280 and 260/230 ratios) were determined using a NanoDrop 1000 device.

Real-time quantitative reverse transcription PCR (RT-qPCR)

Prior to selection of probesets considered for PCR validation, Component and Residual (COR) plots, i.e., plots of the expression of a probeset predicted by the model against the expression of the marker genes with the error added were generated and the linearity of the plots was examined. Variance Inflation Factors (VIFs), that estimate the effect of collinearity on the final fit, were also considered and were required to be <10 for probesets further considered for PCR validation (Fox 2008; Fox and Weisberg 2011). We finally selected for validation two genes predicted by PSEA as differentially expressed in epithelial cells (RORA and TGF-β1) and two in B cells (CERS3 and CAMSAP1). cDNA was transcribed from 75 ng of total RNA utilizing a SuperScript VILO cDNA Synthesis Kit (Invitrogen, USA) in a final volume of 20 µl. The cycle for cDNA synthesis was as follows: 10 min, 25°C; 120 min, 37°C; 5 min 85°C. SYBR-Green-based real-time quantitative PCR (RT-qPCR) was performed using the SYBR™ Select Master Mix (Applied Biosystems, USA) and the CDX96 Real Time PCR Detection System, following a standardized protocol. The thermal cycling conditions were as follows: 2 min, 50°C; 2 min, 95°C; 15 sec 95°C; and 1 min 60°C. Primers were designed for each of the four genes using Primer-BLAST, and

are presented in Supplementary Table 2. mRNA levels were normalized against 18s (internal control) and relative levels were calculated comparing the ΔCt values. One-tailed t-tests, for paired or unpaired observations, as appropriate, based on the availability of pairs of periodontitis-affected/healthy gingival tissue samples with good RNA quality from the same donor, were carried out to test differential expression between gingival health and periodontitis. Statistical significance was defined as p value less than 0.05. Data are presented as mean ± standard error of mean (Figures 2 and 3).

REFERENCES

Almubarak A, Tanagala KKK, Papapanou PN, Lalla E, Momen-Heravi F. 2020. Disruption of monocyte and macrophage homeostasis in periodontitis. Front Immunol. 11:330.

Balfe A, Lennon G, Lavelle A, Docherty NG, Coffey JC, Sheahan K, Winter DC, O'Connell PR. 2018. Isolation and gene expression profiling of intestinal epithelial cells: Crypt isolation by calcium chelation from in vivo samples. Clin Exp Gastroenterol. 11:29-37.

Everitt BS, Hand DJ. 1981. Finite mixture distributions. London: Chapman and Hall.

Fox J. 2008. Applied regression analysis and generalized linear models. Thousand Oaks: SAGE Publications, Inc.

Fox J, Weisberg S. 2011. An r companion to applied regression. Sage.

Gelman A, Hill J. 2006. Data analysis using regression and multilevel/hierarchical models. Cambridge, UK: Cambridge University Press.

Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. 2014. Bayesian data analysis. Boca Raton, Florida, USA: CRC Prress.

Guillaumet-Adkins A, Rodriguez-Esteban G, Mereu E, Mendez-Lago M, Jaitin DA, Villanueva A, Vidal A, Martinez-Marti A, Felip E, Vivancos A et al. 2017. Single-cell transcriptome conservation in cryopreserved cells and tissues. Genome Biol. 18(1):45.

Hasegawa K, Sato A, Tanimura K, Uemasu K, Hamakawa Y, Fuseya Y, Sato S, Muro S, Hirai T. 2017. Fraction of mhcii and epcam expression characterizes distal lung epithelial cells for alveolar type 2 cell isolation. Respir Res. 18(1):150.

Huang L, Yang Y, Yang F, Liu S, Zhu Z, Lei Z, Guo J. 2018. Functions of epcam in physiological processes and diseases (review). International journal of molecular medicine. 42(4):1771-1785.

Hyun SY, Mun S, Kang KJ, Lim JC, Kim SY, Han K, Jang YJ. 2019. Amelogenic transcriptome profiling in ameloblast-like cells derived from adult gingival epithelial cells. Scientific reports. 9(1):3736.

Johnson WE, Li C, Rabinovic A. 2007. Adjusting batch effects in microarray expression data using empirical bayes methods. Biostatistics. 8(1):118-127.

Kebschull M, Guarnieri P, Demmer RT, Boulesteix AL, Pavlidis P, Papapanou PN. 2013. Molecular differences between chronic and aggressive periodontitis. J Dent Res. 92(12):1081-1088.

Kuhn A, Thu D, Waldvogel HJ, Faull RL, Luthi-Carter R. 2011. Population-specific expression analysis (psea) reveals molecular changes in diseased brain. Nat Methods. 8(11):945-947.

McCall MN, Bolstad BM, Irizarry RA. 2010. Frozen robust multiarray analysis (frma). Biostatistics. 11(2):242-253.

McCall MN, Uppal K, Jaffee HA, Zilliox MJ, Irizarry RA. 2011. The gene expression barcode: Leveraging public data repositories to begin cataloging the human and murine transcriptomes. Nucleic Acids Res. 39(Database issue):D1011-1015.

McCall MN, Jaffee HA, Zelisko SJ, Sinha N, Hooiveld G, Irizarry RA, Zilliox MJ. 2014. The gene expression barcode 3.0: Improved data processing and mining tools. Nucleic Acids Res. 42(1):D938-943.

Parmigiani G, Garrett ES, Anbazhagan R, Gabrielson E. 2002. A statistical framework for expression-based molecular classification in cancer. J R Statist Soc B. 64, Part 4:717–736.

Sawle AD, Kebschull M, Demmer RT, Papapanou PN. 2016. Identification of master regulator genes in human periodontitis. J Dent Res. 95(9):1010-1017.

Wu Z, Irizarry RA, Gentleman R, Murillo FM, Spencer F. 2004. A model based background adjustment for oligonucleotide expression. J Am Stat Assoc 2004. 99(4): 909–917.

Wu Z, Irizarry RA. 2005. Stochastic models inspired by hybridization theory for short oligonucleotide arrays. J Comput Biol. 12(6):882-893.

Supplementary Table 1. Marker probesets used in the PSEA analysis

Epithelial cells		Fibroblasts		Endothelial cells		Neutrophils	
Probe id	**Symbol**	**Probe id**	**Symbol**	**Probe id**	**Symbol**	**Probe id**	**Symbol**
1564307_a_at	A2ML1	1555229_a_at	C1S	204677_at	CDH5	206209_s_at	CA4
205623_at	ALDH3A1	231766_s_at	COL12A1	222885_at	EMCN	210789_x_at	CEACAM3
220620_at	C1orf42	231879_at	COL12A1	219436_s_at	EMCN	223552_at	LRRC4
220026_at	CLCA4	212489_at	COL5A1	212951_at	GPR116	207890_s_at	MMP25
224329_s_at	CNFN	202765_s_at	FBN1	203934_at	KDR	1553513_at	VNN3
206642_at	DSG1	221447_s_at	GLT8D2	209087_x_at	MCAM		
219995_s_at	FLJ13841	205422_s_at	ITGBL1	228863_at	PCDH17		
214599_at	IVL	204682_at	LTBP2	221529_s_at	PLVAP		
205470_s_at	KLK11	223690_at	LTBP2	209070_s_at	RGS5		
239381_at	KLK7	212246_at	MCFD2	218353_at	RGS5		
205778_at	KLK7	1557938_s_at	PTRF	206211_at	SELE		
206400_at	LGALS7			204468_s_at	TIE1		
206884_s_at	SCEL						
1554921_a_at	SCEL						
211361_s_at	SERPINB13						
205185_at	SPINK5						
205064_at	SPRR1B						
206008_at	TGM1						
230835_at	UNQ467						
226926_at	ZD52F10						

Monocytes/macrophages		Plasma cells		T cells		B cells	
Probe id	**Symbol**	**Probe id**	**Symbol**	**Probe id**	**Symbol**	**Probe id**	**Symbol**
207270_x_at	CD300C	235965_at	DKFZP434B0335	211861_x_at	CD28	1563469_at	ARID5B
204150_at	STAB1	219910_at	HYPE	206980_s_at	FLT3LG	212715_s_at	MICAL3
38487_at	STAB1	240915_at	IGHV1-69				
		231931_at	PRDM15				

Supplementary Table 2. Primers used in the validation experiments

Gene		
18S	*Forward*	5′-GACCTCATCCCACCTCTCAG-3′
	Reverse	5′-CCATCCAATCGGTAGTAGCG-3′
TGF-β	*Forward*	5’-ACGCAGTACAGCAAGGTCC-3’
	Reverse	5’-GACACAGAGATCCGCAGTCC-3’
RORA	*Forward*	5’-TTGCGGGTGTACCTTGATCC-3’
	Reverse	5’- CTGGCTGCCCCTCAACAATA-3’
CERS3	*Forward*	5’- GGAAGCTTGCTGGAGATTTGC-3’
	Reverse	5’- CAGTACTGGGATGGCAGCAG-3’
CAMSAP1	*Forward*	5’-GAATGATGGCTGCAGTTGGC-3’
	Reverse	5’- GTCATGAGGGTGGGGAATGG-3’

Supplementary Table 3. PSEA -predicted differentially expressed genes by cell type

Epithelial cells

Probe id	Symbol	log2FC	Ref_p	Dif_p	R^2	CC
210479_s_at	RORA	-0.67	1.1E-10	1.1E-07	0.63	1
205637_s_at	SH3GL3	-0.43	2.1E-21	4.5E-07	0.66	1
227309_at	YOD1	-0.45	3.5E-21	1.0E-06	0.62	1
211966_at	COL4A2	-1.59	8.7E-04	3.2E-06	0.77	1
209873_s_at	PKP3	-0.25	1.3E-39	3.4E-06	0.71	0.74
209925_at	OCLN	-0.31	6.5E-30	6.0E-06	0.62	1
203367_at	DUSP14	-0.46	4.2E-16	1.2E-05	0.62	0.88
223544_at	TMEM79	-0.17	2.1E-54	2.6E-05	0.80	1
210059_s_at	MAPK13	-0.27	3.2E-29	3.2E-05	0.63	1
223895_s_at	EPN3	-0.42	1.4E-13	3.3E-05	0.61	1
209203_s_at	BICD2	-0.29	3.0E-26	3.4E-05	0.65	0.86
203430_at	HEBP2	-0.2	3.5E-28	5.9E-04	0.62	0.5
209372_x_at	NA	-0.22	2.1E-29	8.6E-04	0.61	1
216661_x_at	CYP2C9	-0.26	1.5E-21	1.0E-03	0.64	0.6
222190_s_at	C16orf58	-1.46	3.4E-02	1.9E-03	0.64	1
225510_at	OAF	-0.57	2.2E-05	2.7E-03	0.64	1
1553505_at	A2ML1	-0.23	8.6E-25	2.8E-03	0.62	1
219858_s_at	MFSD6	-0.22	7.2E-23	3.2E-03	0.62	1
213533_at	NSG1	-0.26	1.7E-13	4.9E-03	0.65	0.47
226632_at	CYGB	-1.26	3.9E-02	5.7E-03	0.64	1
205464_at	SCNN1B	-0.13	2.4E-38	6.7E-03	0.63	0.5
203085_s_at	TGFB1	-0.69	2.1E-03	0.01	0.63	1
209216_at	WDR45	-0.46	1.0E-04	0.01	0.63	1
227241_at	MUC15	-0.14	5.9E-31	0.02	0.73	1
219476_at	C1orf116	-0.18	2.1E-24	0.02	0.62	0.42

218739_at	ABHD5	-0.17	4.9E-23	0.03	0.65	0.89
218111_s_at	CMAS	-0.17	7.1E-21	0.03	0.66	0.36
203997_at	PTPN3	-0.09	5.5E-37	0.04	0.69	1
1553072_at	BNIPL	-0.13	3.5E-26	0.05	0.61	0.9

Fibroblasts

Probe id	Symbol	log2FC	Ref_p	Dif_p	R^2	CC
201539_s_at	FHL1	0.27	3.7E-26	2.6E-05	0.66	1
211958_at	IGFBP5	0.16	9.5E-23	2.3E-02	0.64	1
201718_s_at	EPB41L2	0.17	2.5E-22	1.6E-02	0.67	0.54
214845_s_at	CALU	0.25	4.3E-14	6.7E-03	0.61	0.39
208851_s_at	THY1	-0.46	3.1E-09	3.5E-03	0.68	1
219315_s_at	TMEM204	-0.71	1.6E-04	6.8E-03	0.73	1
205240_at	GPSM2	-0.68	3.8E-04	0.02	0.60	0.82
208829_at	TAPBP	0.46	8.8E-04	0.02	0.71	0.45
202828_s_at	MMP14	0.57	2.5E-03	0.01	0.64	1
212259_s_at	PBXIP1	0.64	0.01	0.01	0.67	0.5
208872_s_at	REEP5	0.95	0.05	0.001	0.60	1
211633_x_at	IGHG1	0.72	0.05	0.02	0.68	0.25

Endothelial cells

Probe id	Symbol	log2FC	Ref_p	Dif_p	R^2	CC
225369_at	ESAM	-0.41	4.5E-26	5.7E-07	0.76	1
228339_at	ECSCR	-0.44	5.5E-24	3.2E-06	0.68	1
215535_s_at	AGPAT1	-0.65	8.7E-08	1.0E-03	0.65	1
206331_at	CALCRL	0.32	4.9E-14	1.0E-03	0.65	0.22
213131_at	OLFM1	-0.3	7.3E-21	2.0E-03	0.65	0.36

212494_at	TNS2	-0.52	2.1E-07	6.0E-03	0.65	1
201389_at	ITGA5	-0.29	3.0E-15	0.01	0.68	1
212902_at	SEC24A	0.75	0.01	0.01	0.63	0.15
209166_s_at	MAN2B1	-0.68	9.73E-05	0.01	0.67	1
205247_at	NOTCH4	-0.3	1.08E-13	0.02	0.63	1
209474_s_at	ENTPD1	0.5	4.12E-04	0.02	0.68	0.71
1552256_a_at	SCARB1	-0.65	9.11E-05	0.02	0.66	0.45
200827_at	PLOD1	-0.35	5.33E-09	0.04	0.71	1

Neutrophils

Probe id	Symbol	log2FC	Ref_p	Dif_p	R^2	CC
226907_at	PPP1R14C	-1.14	5.7E-09	2.4E-10	0.63	1
223694_at	TRIM7	-0.99	1.1E-07	1.9E-08	0.62	1
202428_x_at	DBI	-3.19	1.3E-02	1.3E-05	0.81	1
232116_at	GRHL3	-0.74	2.7E-07	2.5E-05	0.75	1
227736_at	C10orf99	-1.15	2.1E-04	4.5E-05	0.62	1
220013_at	EPHX3	-0.95	3.2E-05	5.4E-05	0.67	1
204203_at	CEBPG	-0.39	5.2E-14	6.3E-05	0.60	1
230769_at	DENND2C	-0.97	7.9E-05	1.1E-04	0.71	1
214626_s_at	GANAB	0.83	1.0E-02	1.1E-04	0.75	1
228587_at	FAM83G	-0.43	4.8E-12	2.2E-04	0.63	1
204616_at	UCHL3	-0.61	3.1E-07	3.1E-04	0.66	1
209311_at	BCL2L2	-1.2	3.1E-03	4.7E-04	0.60	1
216025_x_at	CYP2C9	-0.77	6.5E-05	7.9E-04	0.63	0.58
201315_x_at	IFITM2	0.75	2.0E-02	1.2E-03	0.70	1
209569_x_at	NSG1	-0.71	1.3E-04	2.6E-03	0.66	0.8
221854_at	PKP1	-0.85	4.2E-03	7.6E-03	0.67	0.5
212702_s_at	BICD2	-0.46	2.1E-05	0.01	0.61	0.93

224615_x_at	HM13	0.55	1.0E-02	0.02	0.68	1
209880_s_at	SELPLG	0.28	2.7E-06	0.02	0.65	0.56
218084_x_at	FXYD5	-0.85	2.0E-02	0.04	0.72	0.5
200770_s_at	LAMC1	0.34	2.2E-04	0.04	0.80	0.6

Monocytes/Macrophages

Probe id	Symbol	log2FC	Ref_p	Dif_p	R^2	CC
210657_s_at	SEPT4	0.75	5.20E-03	1.30E-04	0.66	0.67
220532_s_at	TMEM176B	0.35	1.90E-08	1.10E-03	0.72	0.57
202112_at	VWF	0.26	1.40E-08	0.01	0.73	0.5
204503_at	EVPL	-0.32	2.40E-08	0.02	0.60	0.55
211881_x_at	IGLJ3	0.38	5.20E-05	0.02	0.73	0.18

Plasma cells

Probe id	Symbol	log2FC	Ref_p	Dif_p	R^2	CC
212890_at	SLC38A10	-1.17	1.3E-11	1.4E-05	0.76	1
55093_at	CHPF2	-1.3	1.4E-09	4.2E-05	0.70	1
206593_s_at	MED22	-2.75	1.5E-03	2.6E-03	0.61	1
202908_at	WFS1	-0.72	4.3E-11	3.1E-03	0.71	0.17
223065_s_at	STARD3NL	-1.2	5.8E-06	3.7E-03	0.61	0.52
201206_s_at	RRBP1	-0.99	2.4E-05	0.02	0.65	0.67
204158_s_at	TCIRG1	-1.21	2.5E-04	0.02	0.73	1
217861_s_at	PREB	-0.71	1.7E-07	0.02	0.69	0.58
204683_at	ICAM2	0.81	5.0E-03	0.02	0.78	0.80
200644_at	MARCKSL1	-1.49	3.9E-03	0.04	0.66	1
202369_s_at	TRAM2	-0.55	3.7E-08	0.05	0.69	1

T cells

Probe id	Symbol	log2FC	Ref_p	Dif_p	R^2	CC
224252_s_at	FXYD5	-0.34	5.9E-03	0.02	0.17	0.17
215346_at	CD40	0.28	0.02	0.03	0.25	0.25
209496_at	RARRES2	0.34	0.07	0.04	0.25	0.25

B cells

Probe id	Symbol	log2FC	Ref_p	Dif_p	R^2	CC
222538_s_at	APPL1	-2.58	8.8E-05	4.E-06	0.61	0.89
202539_s_at	HMGCR	-0.89	1.2E-09	4.E-05	0.77	1
204552_at	INPP4A	0.86	0.01	3.E-03	0.65	1
200971_s_at	SERP1	0.74	4.0E-03	5.E-03	0.66	0.56
204678_s_at	KCNK1	-0.64	4.1E-07	5.E-03	0.61	0.88
204912_at	IL10RA	0.6	2.2E-04	6.E-03	0.67	1
210785_s_at	THEMIS2	0.85	0.02	6.E-03	0.68	1
212712_at	CAMSAP1	-0.89	9.33E-05	8.E-03	0.64	1
220306_at	FAM46C	0.83	0.02	0.01	0.70	1
212324_s_at	VPS13D	-1	0	0.02	0.65	0.89
200866_s_at	PSAP	0.71	0.01	0.02	0.64	0.28
1554252_a_at	CERS3	-1.65	0.02	0.02	0.73	1
206896_s_at	GNG7	0.78	0.04	0.04	0.62	1

Supplemental Figure 1. Flowchart of the PSEA computational steps