Supplemental Data File 1

Cross-disease transcriptomics: Unique IL-17A signaling in psoriasis lesions and an autoimmune PBMC signature

William R. Swindell[1,2§], Mrinal K. Sarkar[2], Yun Liang[2], Xianying Xing[2], Johann E. Gudjonsson[2]

[1]Ohio University Heritage College of Osteopathic Medicine, Athens, OH, 45701 USA

[2]University of Michigan, Department of Dermatology, Ann Arbor, MI, 48109-2200 USA

[§]Corresponding Author

Email addresses:

WRS: ws277814@ohio.edu

MKS: mksarkar@ med.umich.edu

YL: yunlia@ med.umich.edu

XX: xingx@ med.umich.edu

JEG: johanng@med.umich.edu

Supplemental Methods

RNA-seq Datasets

RNA-seq meta-analysis was performed using reads from four studies utilizing the Illumina Genome Analyzer platform (GSE41745, GSE54456/GSE63979 and GSE67785) 44 patients) (Jabbari *et al.*, 2012; Li *et al.*, 2014; Swindell *et al.*, 2015; Tsoi *et al.*, 2015). Reads in each study had been generated from lesional (PP) and uninvolved (PN) skin of psoriasis patients. Skin samples were obtained from consenting patients following a 2-4 week washout period during which patients did not use systemic medications. PN samples were macroscopically normal skin obtained from the buttock/upper thigh (GSE54456/GSE63979 and GSE67785) (Li *et al.*, 2014; Swindell *et al.*, 2015; Tsoi *et al.*, 2015) or a region anatomically comparable to the PP biopsy site in a given patient (GSE41745) (Jabbari *et al.*, 2012). The complete meta-dataset included RNA-seq reads from 88 paired PP and PN samples (n = 44 patients), based upon reads from GSE41745 (n = 3), GSE54456/GSE63979 (n = 27) and GSE6778 (n = 14). Informed written consent was obtained in each study from volunteer patients in accordance with Declaration of Helsinki principles, with protocols approved by institutional review boards of participating institutions (University of Michigan, Ann Arbor, MI; Rockefeller University, New York, NY).

RNA-seq read processing and gene expression quantification

Sequence files in fastq format were downloaded from the Sequence Read Archive using SRA Toolkit version 2.3.5 (function: fastq-dump) (Kodama *et al.*, 2012). Quality scores were encoded in phred33 format, or otherwise, Trimmomatic was used to convert native scores to phred33 format (option: TOPHRED33) (Bolger *et al.*, 2014). This ensured that all files were

homogenous with regard to quality score encoding. FastQC was used to assess sequence quality before and after read processing (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Read processing steps included adaptor removal (cutadapt), read trimming (fastx-toolkit), running-sum filtering (cutadapt), and window-based filtering (fastx-toolkit) (http://hannonlab.cshl.edu/fastx_toolkit/; Martin, 2011). The initial 10 bases were removed from reads, since FastQC analysis had indicated nucleotide frequency bias at these bases for each dataset, as previously observed for reads generated using the Illumina platform (Hansen *et al.*, 2010). The cutadapt running-sum filter was then applied to remove reads with low phred33 scores (with –q 25 argument) (Martin, 2011). The fastx-toolkit window-based filter was next used to remove reads with phred33 scores less than 25 at > 50% of read sequence (i.e., arguments -q 25 -p 50) (http://hannonlab.cshl.edu/fastx_toolkit/). Following these preprocessing steps, the average number of reads per sample was reduced from 30.2 to 25.1 million (GSE41745), from 45.2 to 31.6 million (GSE54456/GSE63979), and from 29.3 to 29.0 million (GSE67785).

Processed reads were mapped to the UCSC hg19 genome using tophat (version: 2.0.13) (Kim *et al.*, 2013). Reads were mapped using UCSC hg19 transcript annotation files in GTF format (-G option) with coverage based search for exon junctions disabled (--no-coverage-search). The average percentage of mapped reads among samples was 83.8% (GSE41745), 93.0% (GSE54456/GSE63979) and 95.1% (GSE67785). Average intragenic mapping rates were 80.7% (GSE41745), 91.0% (GSE54456/GSE63979) and 92.7% (GSE67785). Bam files generated by tophat were sorted and indexed using SAMtools (Li *et al.*, 2009). The number of reads mapping to each gene feature was then tabulated using HTSeq (function: htseq-count) (Anders *et al.*, 2015). Reads with alignment quality lower than 10 were excluded from

tabulations (-a 10). Reads were also excluded if they mapped ambiguously to multiple features or if they partially mapped to sequence outside of known exons (-m intersection-strict). Cufflinks was used to quantify expression in terms of the number of fragments per kilobase of transcript per million mapped reads (FPKM) (Trapnell *et al.*, 2012). Read mapping rates and other quality control indices for each sample were calculated using RSeQC and RNA-SeQC (DeLuca *et al.*, 2012; Wang *et al.*, 2012).

RNA-seq differential expression analysis

Differential expression analysis was performed for 15643 protein-coding features with detectable expression in at least 25% (22) of the 88 PP and PN samples. A feature was considered detected in a sample if the lower FPKM 95% confidence limit generated by Cufflinks was greater than zero, and if the count per million mapped reads (cpm) was greater than 0.25 (Swindell *et al.*, 2014). To evaluate differential expression, raw counts from each dataset were first normalized independently using the voom algorithm, which generates $\log_2$-normalized expression values appropriate for differential expression tests relying on the normal distribution (Law *et al.*, 2014). Voom-normalized expression values for each dataset were generated using patient as a blocking variable while utilizing inter-patient correlation estimates (i.e., voom options "block" and "correlation"). PP versus PN differences in voom-normalized expression estimates were calculated for each patient, and these differences were then pooled across datasets to yield a meta-dataset with differences from 44 patients. Significant changes in gene expression were evaluated by linear modeling with empirical Bayes-based shrinkage of gene-wise variance estimates towards a common value (R package: limma; functions: lmFit and eBayes) (Smyth, 2004). To control the false discovery rate (FDR), raw p-values generated from limma linear

models were adjusted using the Benjamini-Hochberg method (Benjamini and Hochberg, 1995). Genes with two-fold change in expression (PP/PN > 2.0 or < 0.50) and FDR-adjusted p-value less than 0.05 were considered differentially expressed (DEGs).

Additional Data Resources

A microarray gene signature database was used to compare our DEG list with changes in gene expression observed in other human diseases, cultured KCs, and PBMCs (Swindell *et al.*, 2015). The complete database includes gene signatures from 2178 treatment comparisons associated with 21,337 unique Gene Expression Omnibus (GEO) samples (Swindell *et al.*, 2015). A subset of this database was used in the current study, including gene signatures associated with other human skin diseases (98 signatures), KC cytokine responses (59 signatures), and PBMC tissue (421 signatures) (see Supplemental Data File 2). The 98 human skin disease gene signatures were generated from microarray comparisons between diseased and macroscopically normal skin samples (full thickness biopsies). The 59 KC cytokine signatures were generated from comparisons between cytokine-treated KCs and control cells (untreated or mock-treated). The 421 PBMC signatures were generated from comparisons between groups of patients that differed according to disease status, phenotypic traits, or treatments (e.g., pre- and post-vaccination). For each signature and treatment comparison, differential expression statistics were calculated for all genes represented on the microarray platform used by a given study (i.e., estimated fold-change, raw p-value and FDR-adjusted p-value) (Swindell *et al.*, 2015). Linear modeling with moderated t-statistics (limma) was used to assess differential expression (see above) (Smyth, 2004). To ensure that differences in sample or microarray processing between

laboratories and studies did not influence our results, in all cases comparisons were only made between samples derived from the same GEO series accession.

Psoriasis Specificity Index (PSI)

Psoriasis-specific and non-specific DEGs were distinguished using a psoriasis specificity index (PSI). The PSI was calculated based upon gene expression changes in n = 98 microarray treatment comparisons between diseased and macroscopically normal skin. For the jth comparison, changes in DEG expression were quantified based upon the p-value (P_j) generated from the two sample linear model comparison of expression in diseased and normal skin. The p-value was $\log_{10}$-transformed and multiplied by a directionality indicator (d_j) defined as -1 if the change in gene expression was psoriasis-consistent and 1 if the change was psoriasis-inconsistent. This calculation was performed for all treatment comparisons and PSI was defined as the average value across the n comparisons (Equation 1).

$$[1] \qquad PSI = \frac{1}{n}\left(\sum_{j=1}^{n} -d_j \log_{10} P_j\right)$$

PSI thus reflects both the directionality and magnitude of psoriasis DEG expression changes in other human skin diseases, allowing discrimination between DEGs that are psoriasis-specific (PSI ≈ 0 or PSI > 0) and non-specific (PSI < 0). A PSI of zero indicates that the DEG is not consistently altered in other skin diseases. A value of -1.30 or less indicates that, on average, the DEG shows significant ($P < 0.05$) psoriasis-like gene expression changes in other diseases. Alternatively, a value of 1.3 or greater would indicate that, on average, the DEG shows significant ($P < 0.05$) gene expression changes in other skin diseases that are opposite in direction to those observed in psoriasis lesions.

Most microarray platforms provide only partial genome coverage and therefore some psoriasis DEGs identified by RNA-seq are not represented by probes. Of 207074 p-values used for PSI calculations (2113 DEGs × 98 microarray treatment comparisons), 74.4% (154063) could be estimated directly from microarray expression values, but the remaining 25.6% (53011) were missing due to lack of probe representation on microarray platforms. For these 53011 instances, p-values required by Equation 1 (above) were obtained using nearest neighbor imputation, with nearest neighbors of psoriasis DEGs identified by co-expression analysis (using an independent RNA-seq dataset; described below). Given a missing DEG p-value for experiment *j*, we identified p-values for the top 10 non-missing genes most strongly co-expressed with that DEG, and the average of these 10 p-values was calculated and imputed into Equation 1. Thus, for DEGs not represented in a given microarray experiment, treatment effects (p-values) were approximated based upon p-values obtained for the 10 most strongly co-expressed genes measured in the experiment.

An independent RNA-seq dataset was needed to identify nearest neighbors most strongly co-expressed with psoriasis DEGs. For this purpose, we utilized 162 skin biopsies analyzed by RNA-seq, which included 72 PP samples and 90 normal skin samples from control subjects (GSE54456/GSE63979) (Li *et al.*, 2014; Tsoi *et al.*, 2015). These samples had not been utilized for identification of psoriasis DEGs (above), since we only used patients with paired PP and PN samples to identify DEGs. The 162 samples were processed and mapped to the UCSC hg19 genome following the same RNA-seq processing steps outlined above. Gene read counts were calculated for each sample (HTSeq) and these were subsequently normalized using the voom algorithm (Law *et al.*, 2014). A gene distance matrix was then calculated based upon Spearman correlation coefficients ($1 - r_s$) between expression profiles (n = 162 samples). This matrix was

then used to identify genes most strongly co-expressed with psoriasis DEGs, and these nearest neighbor genes were then used if necessary to obtain an imputation p-value for Equation (1).

Cytokine-treated keratinocytes

Normal human keratinocyte (NHK) cultures were established in serum-free medium (Medium 154, Invitrogen, Portland, OR). Three independent cell lines from separate donors were treated identically and used as biological replicates. Second or third-passage cells from each lineage were maintained 4 days post-confluence and starved of growth factors for 24 hours (0.1 mM calcium). Cells were subsequently treated with recombinant human IL-17A (20 ng/mL; R&D Systems, Minneapolis, MN) for 8 or 24 hours prior to RNA extraction.

Immunohistochemistry

Lesional (PP) and uninvolved (PN) tissue samples were obtained following protocols described previously (Swindell *et al.*, 2013). Informed written consent was obtained from all patients using procedures approved by the University of Michigan institutional review board (HUM00037994). Paraffin embedded tissue sections (PP, PN and NN) were heated at 60^{o}C for 30 minutes, de-paraffinized, and rehydrated. Slides were placed in PH9 antigen retrieval buffer and heated at 95^{o}C for 20 minutes in a hot water bath. After cooling, slides were treated with 3% H_2O_2 (5 minutes) and blocked using 10% goat serum (30 minutes). Overnight incubation (4^{o}C) was then performed using anti-human ATP1B1 (Lifespan Biosciences, cat. no. LS-B5781/68401) at a concentration of 1µg/ml. Slides were then washed, treated with secondary antibody, peroxidase (30 minutes) and diaminobenzidine substrate.

Real time quantitative reverse transcription PCR (RT-PCR)

Lesional and uninvolved skin biopsies were flash-frozen and a rotor-stator homogenizer was used to disrupt and homogenize samples. RNA was extracted using the RNeasy Fibrous Tissue kit (Qiagen, cat. no. 74704) with on-column DNAase digestion (Qiagen, cat. no. 79254). RNA from cultured KCs was isolated using the RNeasy Protect Cell Mini Kit (Qiagen, cat. no. 74624) with on-column DNAase digestion (Qiagen, cat. no. 79254). RNA quantification was performed using a NanoDrop spectrophotometer and the Agilent Bioanalyzer was used to assess RNA quality. Pre-designed primer assays were used for amplification of *ATP1B1* (Applied Biosystems, cat. no. Hs00426868_g1), *PON2* (Hs00165563_m1), *KRT16* (Hs04194235_g1) and *STAT1* (Hs01013996_m1). Cycle threshold (Ct) values were normalized using *RPLP0* (Hs99999902_m1) as an endogenous control gene.

Supplemental References

Anders S, Pyl PT, Huber W (2015) HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics (Oxford, England)* 31:166-9.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a powerful and practical approach to multiple testing. *J Roy Stat Soc B* 57:289–300.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)* 30:2114-20.

DeLuca DS, Levin JZ, Sivachenko A, *et al.* (2012) RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics (Oxford, England)* 28:1530-2.

Hansen KD, Brenner SE, Dudoit S (2010) Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic acids research* 38:e131.

Jabbari A, Suarez-Farinas M, Dewell S, *et al.* (2012) Transcriptional profiling of psoriasis using RNA-seq reveals previously unidentified differentially expressed genes. *The Journal of investigative dermatology* 132:246-9.

Kim D, Pertea G, Trapnell C, *et al.* (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome biology* 14:R36.

Kodama Y, Shumway M, Leinonen R (2012) The Sequence Read Archive: explosive growth of sequencing data. *Nucleic acids research* 40:D54-6.

Law CW, Chen Y, Shi W, *et al.* (2014) voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome biology* 15:R29.

Li B, Tsoi LC, Swindell WR, *et al.* (2014) Transcriptome analysis of psoriasis in a large case-control sample: RNA-seq provides insights into disease mechanisms. *The Journal of investigative dermatology* 134:1828-38.

Li H, Handsaker B, Wysoker A, *et al.* (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)* 25:2078-9.

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* 17:10-2.

Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical applications in genetics and molecular biology* 3:Article3.

Swindell WR, Johnston A, Xing X, *et al.* (2013) Robust shifts in S100a9 expression with aging: a novel mechanism for chronic inflammation. *Scientific reports* 3:1215.

Swindell WR, Remmer HA, Sarkar MK, *et al.* (2015) Proteogenomic analysis of psoriasis reveals discordant and concordant changes in mRNA and protein abundance. *Genome medicine* 7:86.

Swindell WR, Xing X, Voorhees JJ, *et al.* (2014) Integrative RNA-seq and microarray data analysis reveals GC content and gene length biases in the psoriasis transcriptome. *Physiological genomics* 46:533-46.

Trapnell C, Roberts A, Goff L, *et al.* (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols* 7:562-78.

Tsoi LC, Iyer MK, Stuart PE, *et al.* (2015) Analysis of long non-coding RNAs highlights tissue-specific expression patterns and epigenetic profiles in normal and psoriatic skin. *Genome biology* 16:24.

Wang L, Wang S, Li W (2012) RSeQC: quality control of RNA-seq experiments. *Bioinformatics (Oxford, England)* 28:2184-5.

Supplemental Figures

Figure S1. RNA-seq data from different laboratories diverge in their expression profiles but are similar with regard to expression differences between lesional (PP) and uninvolved (PN) skin. (A) PP and PN samples were plotted with respect to the first two principal component axes (*n* = 88 samples total). (B) DEG scores were calculated based upon 100 PP-increased DEGs (horizontal axis) and 100 PP-decreased DEGs (vertical axis) previously identified from microarray meta-analysis (*n* = 237 patients). For each patient, scores represent the average RNA-seq-based $\log_2$-transformed FC estimate (PP/PN) among the 100 PP-increased microarray-based DEGs (horizontal axis) and 100 PP-decreased microarray-based DEGs (vertical axis). (C) The 44 patients were clustered (complete linkage) based upon genome-wide expression differences between PP and PN samples (Euclidean distance).

Figure S2. Meta-analysis improves correlation between RNA-seq and microarray fold-change estimates. Fold-change estimates (PP/PN) from a previous microarray meta-analysis (*n* = 237 patients) were compared to those obtained using RNA-seq. The microarray estimates were compared to those calculated from individual RNA-seq studies (A – C), as well as those calculated from meta-analysis of the three RNA-seq studies (D). The spearman correlation coefficient estimate is shown in each plot (lower right). Yellow ellipses outline the 80% of genes closest to the bivariate mean (Mahalanobis distance).

Figure S3. Overlap among psoriasis DEGs identified from individual RNA-seq datasets and meta-analysis. Psoriasis DEGs were identified based upon PP and PN samples from three individual RNA-seq datasets (GSE41745, *n* = 3 patients; GSE54456/GSE63979, *n* = 27;

GSE67785, $n = 14$) and a pooled meta-analysis of samples from all three datasets ($n = 44$). RNA-seq reads from each source were processed using the same methodology and DEGs were identified using the same thresholds (FDR < 0.05 with FC > 2 or FC < 0.50). Venn diagrams show the overlap among (A) PP-increased and (B) PP-decreased DEGs.

Figure S4. Genes with decreased expression in psoriasis lesions are longer and weakly expressed (RNA-seq meta-analysis). RNA-seq meta-analysis was used to identify 2113 DEGs in lesional psoriasis skin (FC > 2 or < 0.50 with FDR < 0.05). Genes were binned according to (A) gene length, (B) GC content, or (C) average FPKM (PP and PN skin) and the percentage of DEGs within each bin is shown. The percentage of PP-increased and PP-decreased DEGs within each bin is indicated (red: PP-increased; blue: PP-decreased; yellow font: FDR < 0.05, Fisher's Exact Test). Asterisk symbols denote a significantly increased percentage of DEGs within a given bin (PP-increased + PP-decreased; FDR < 0.05, Fisher's Exact Test). (D) Overlap between DEGs identified from RNA-seq ($n = 44$ patients) and microarray ($n = 237$) meta-analyses. (E) Mean FPKM among DEGs identified specifically by RNA-seq and microarray. Mean FPKM was calculated across 88 PP and PN samples and the median among DEGs (middle line: median; boxes span lower and upper quartiles). (F) Correlation and RNA-seq and microarray FC estimates (PP/PN) among genes with different expression levels (FPKM). Correlations are shown for each dataset individually ($3 \leq n \leq 27$) and the meta-dataset ($n = 44$).

Figure S5. Genes with decreased expression in psoriasis lesions are longer and weakly expressed (Analysis of individual datasets). Analyses shown in Figure S4A – S4C were repeated with respect to each individual RNA-seq dataset (GSE41745, GSE54456/GSE63979

and GSE67785). Genes were binned according to gene length (A – C), GC content (D – F), or average FPKM in PP and PN skin (G – I). The percentage of DEGs within each bin is shown along with the separate contributions of PP-increased DEGs (red) and PP-decreased DEGs (blue) (yellow font: FDR < 0.05, Fisher's Exact Test). Asterisk symbols denote a significantly increased percentage of DEGs within a given bin (PP-increased + PP-decreased; FDR < 0.05, Fisher's Exact Test).

Figure S6. Most psoriasis cutaneous DEGs are similarly altered in other human skin diseases (psoriasis specificity index). Psoriasis specificity index (PSI) was calculated for (A) 954 PP-increased DEGs and (B) 1159 PP-decreased DEGs. Most DEGs are non-specific as indicated by negative PSI values (right). Heatmaps show changes in DEG expression in 98 treatment comparisons between diseased and normal skin. DEG expression changes are quantified using signed $\log_{10}$-transformed p-values (Log_{10}P; positive values: increased expression in diseased sin; negative values: decreased expression in diseased skin). In both (A) and (B), rows and columns have been clustered using hierarchical clustering (average linkage) and the Euclidean distance. Disease classifications are indicated above each heatmap (see legend).

Figure S7. The most and least psoriasis-specific DEGs and their changes in expression across human skin disease studies. (A) PP-increased DEGs. PSI was used to identify the 15 least psoriasis-specific PP-increased DEGs (top) and the 15 most psoriasis-specific PP-increased DEGs (bottom). (B) PP-decreased DEGs. PSI was used to identify the 15 least psoriasis-specific PP-decreased DEGs (top) and the 15 most psoriasis-specific PP-decreased DEGs (bottom). In

both (A) and (B), DEGs are listed in the left margin and the estimated PSI is shown in parentheses.

Figure S8. Gene Ontology (GO) Biological process (BP) terms with differential enrichment between psoriasis-specific and non-specific DEGs. (A) GO BP terms enriched with respect to psoriasis-specific but not non-specific PP-increased DEGs. (B) GO BP terms enriched with respect to non-specific but not psoriasis-specific PP-increased DEGs. (C) GO BP terms enriched with respect to psoriasis-specific but not non-specific PP-decreased DEGs. (D) GO BP terms enriched with respect to non-specific but not psoriasis-specific PP-decreased DEGs. In (A) – (D), enrichment of GO BP terms was compared between the 100 most psoriasis-specific DEGs (highest PSI) and 100 least psoriasis-specific DEGs (lowest PSI). Horizontal axes show $-\log_{10}$-transformed p-values obtained from testing whether the indicated GO BP term is enriched with respect to each gene set. Example DEGs associated with each GO BP term are listed in yellow font.

Figure S9. PP-increased DEGs (psoriasis lesions) and PP-decreased DEGs (PBMC) overlap significantly with psoriasis GWAS signals. Overlap was evaluated between genes surrounding psoriasis-associated SNPs and (A, B) cutaneous PP-increased DEGs, (C, D) cutaneous PP-decreased DEGs, (E) DEGs with decreased expression in psoriasis patient PBMC, and (F) DEGs with decreased expression in psoriasis and multiple sclerosis (MS) patient PBMC. Each figure compares the GWAS/DEG and GWAS/non-DEG overlap with respect to varying window sizes (0 – 200 kb; horizontal axis), where window size reflects the amount of sequence scanned to identify genes near psoriasis-associated SNPs. Larger window sizes are more

permissive and expand the number of genes associated with GWAS signals. For each window size, we tested whether GWAS/DEG overlap was significantly greater than GWAS/non-DEG overlap (yellow: $P < 0.05$, Fisher's Exact Test).

Figure S1

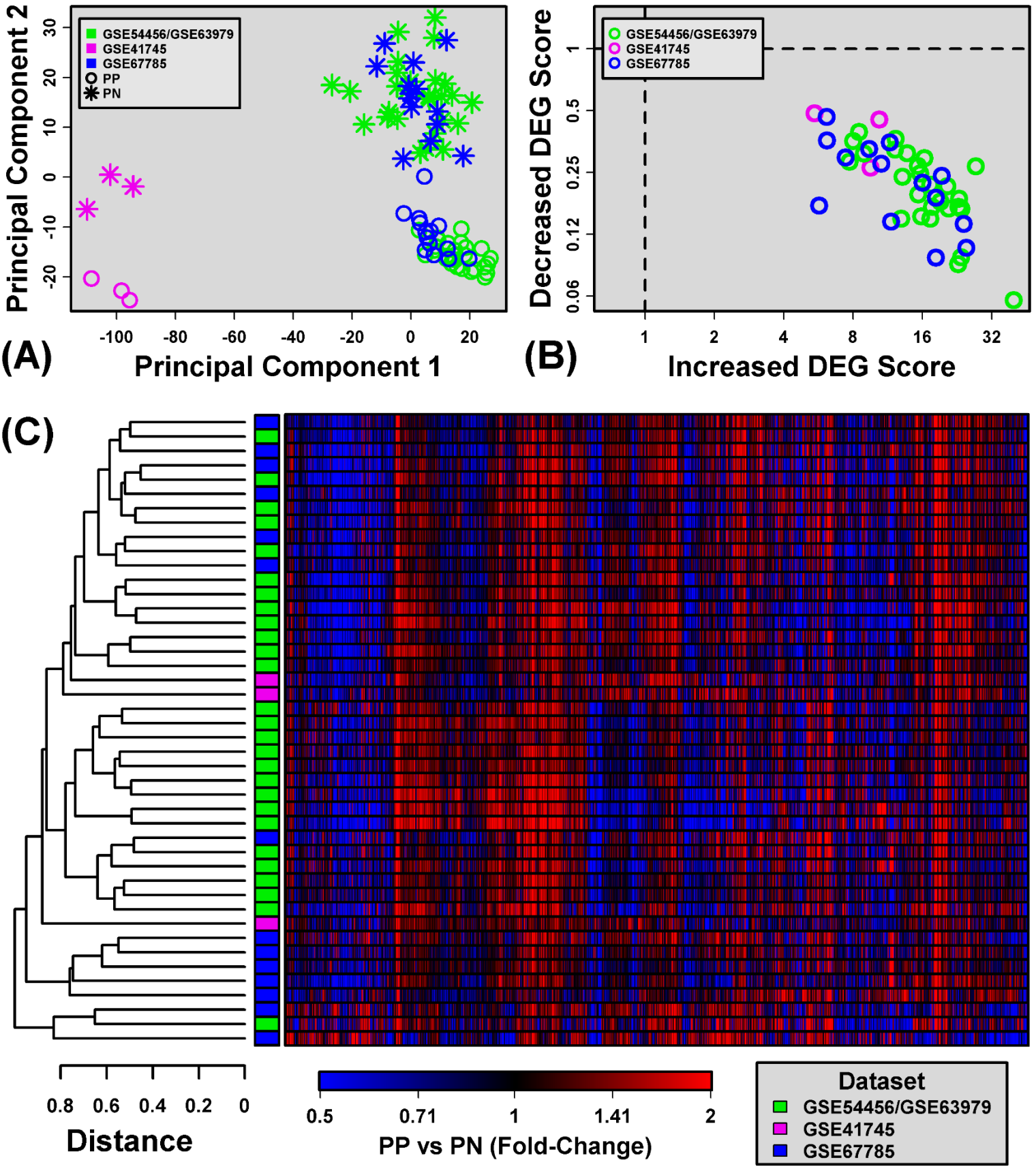

GSE54456/GSE63979
GSE41745
GSE67785
PP
PN
Principal Component 2
Principal Component 1
(A)
Decreased DEG Score
Increased DEG Score
(B)
(C)
Distance
PP vs PN (Fold-Change)
Dataset
GSE54456/GSE63979
GSE41745
GSE67785

Figure S2

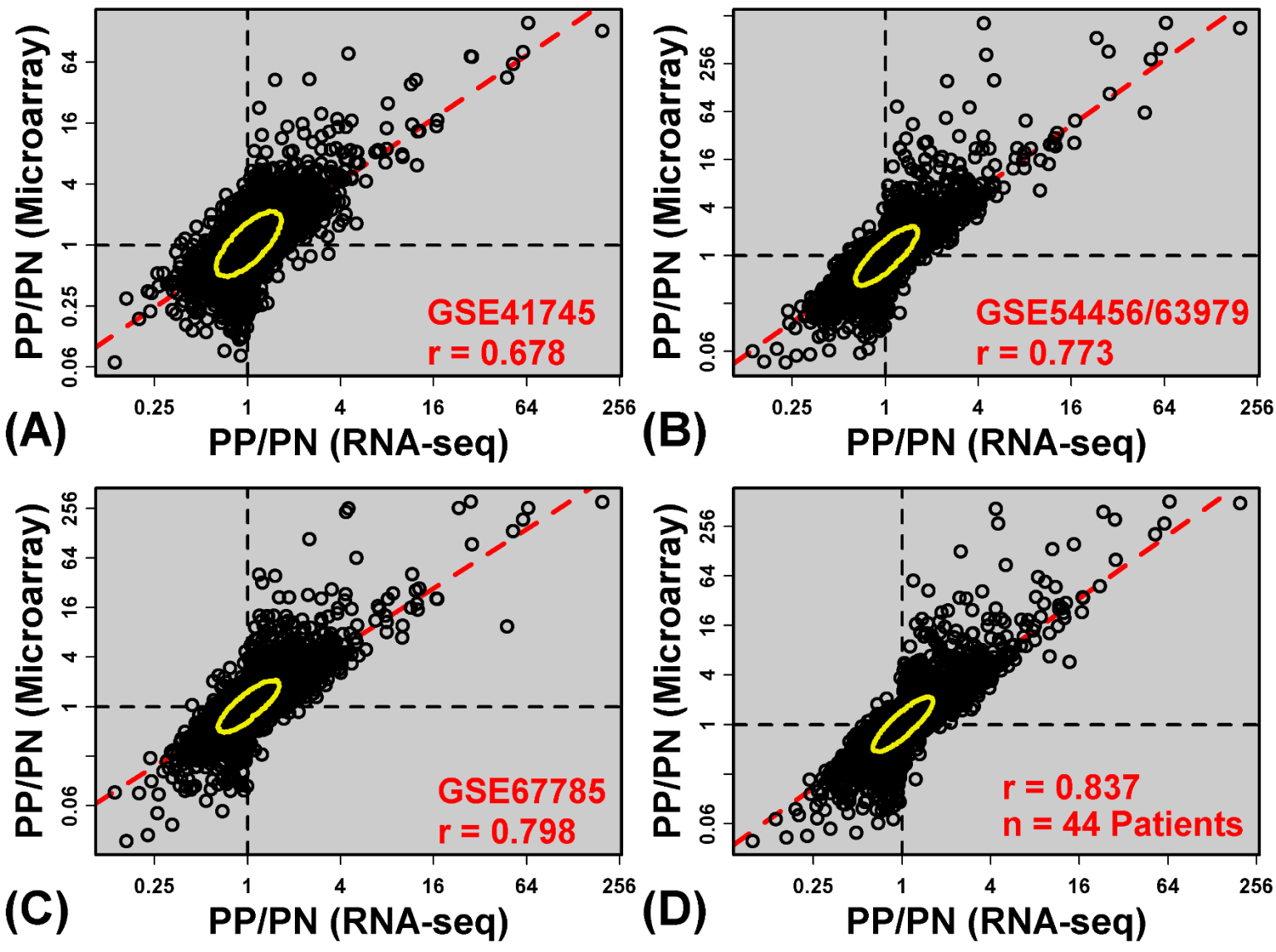
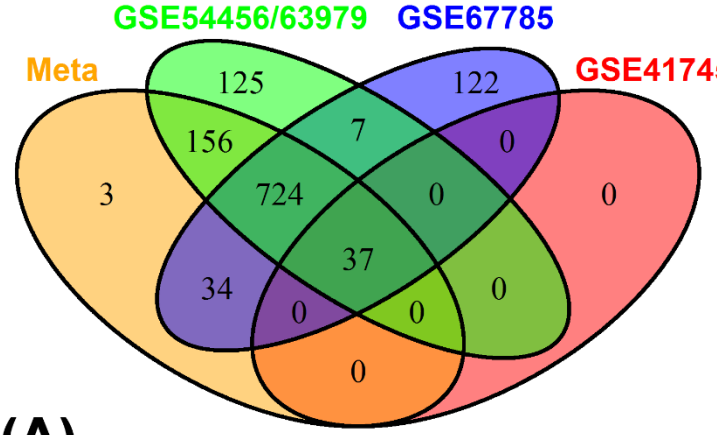

PP/PN (Microarray)
PP/PN (RNA-seq)
0.06
0.25
1
4
16
64
256
GSE41745
r = 0.678
(A)
GSE54456/63979
r = 0.773
(B)
GSE67785
r = 0.798
(C)
r = 0.837
n = 44 Patients
(D)

Figure S3


GSE54456/63979
GSE67785
Meta
GSE41745
125
122
156
7
0
3
724
0
0
37
34
0
0
0
0
(A)

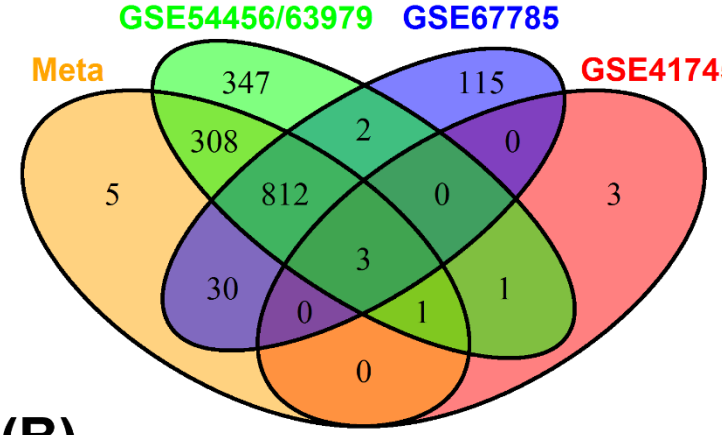
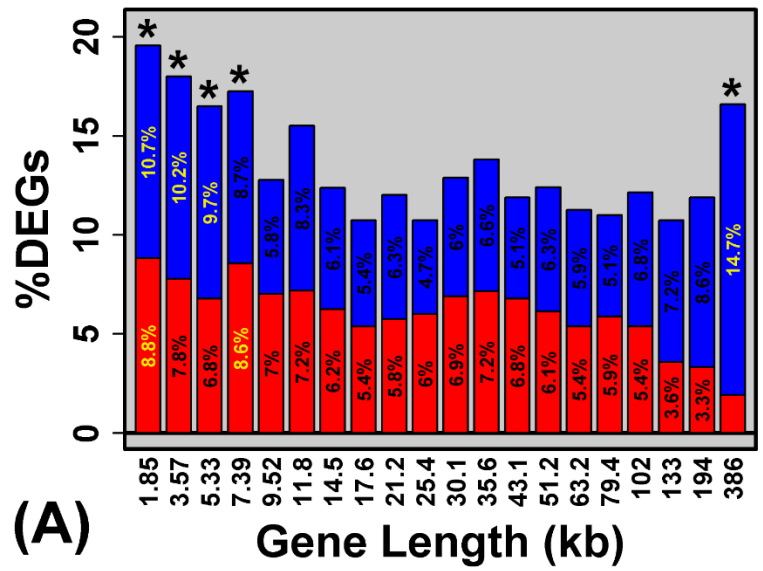
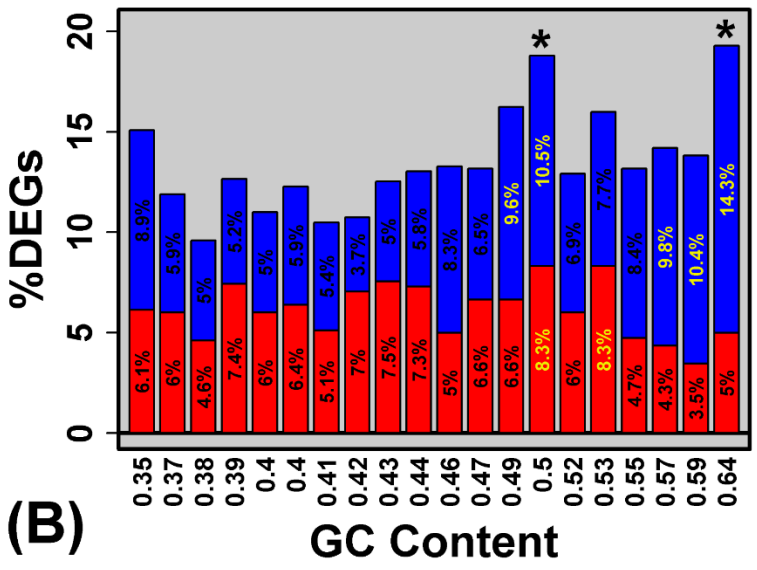

GSE54456/63979
GSE67785
Meta
GSE41745
347
115
308
2
0
5
812
0
3
3
30
0
1
1
0
(B)

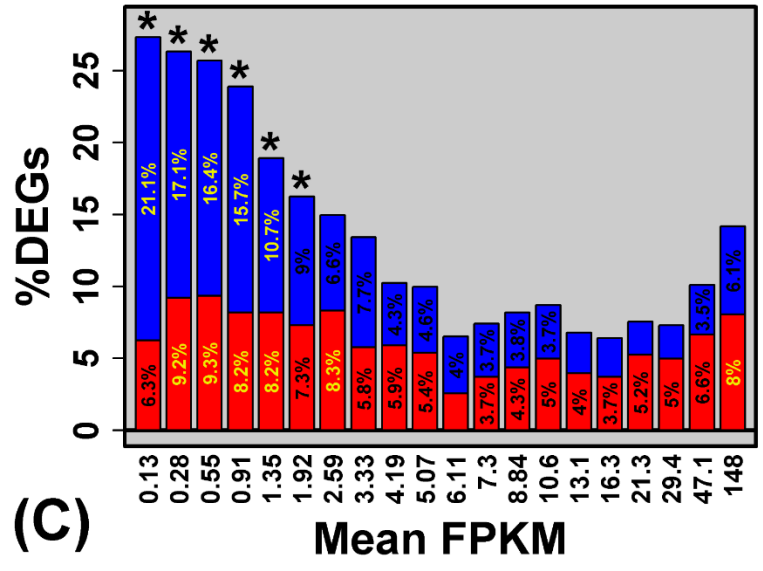
Figure S4



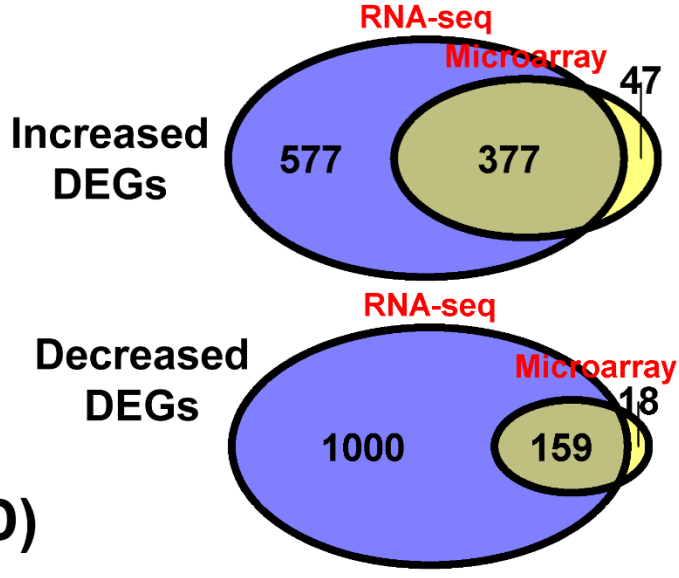
(A)



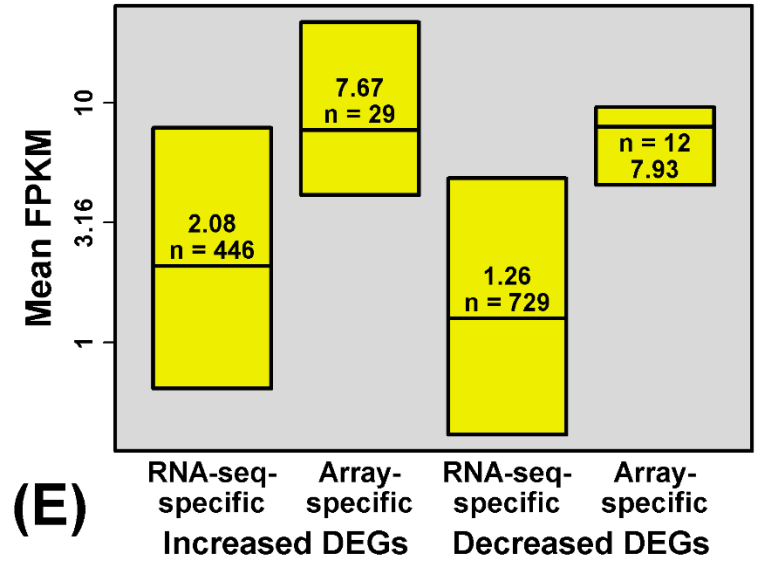
(B)



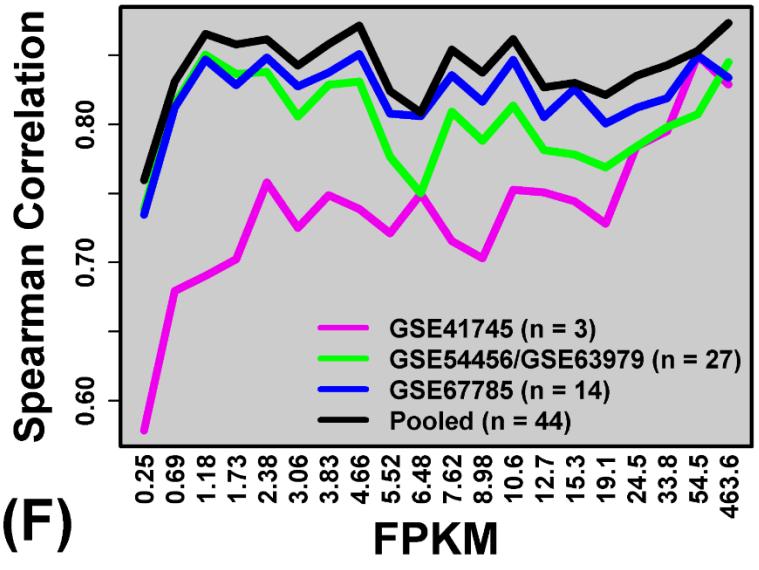
(C)



(D)



(E)



(F)

Figure S5

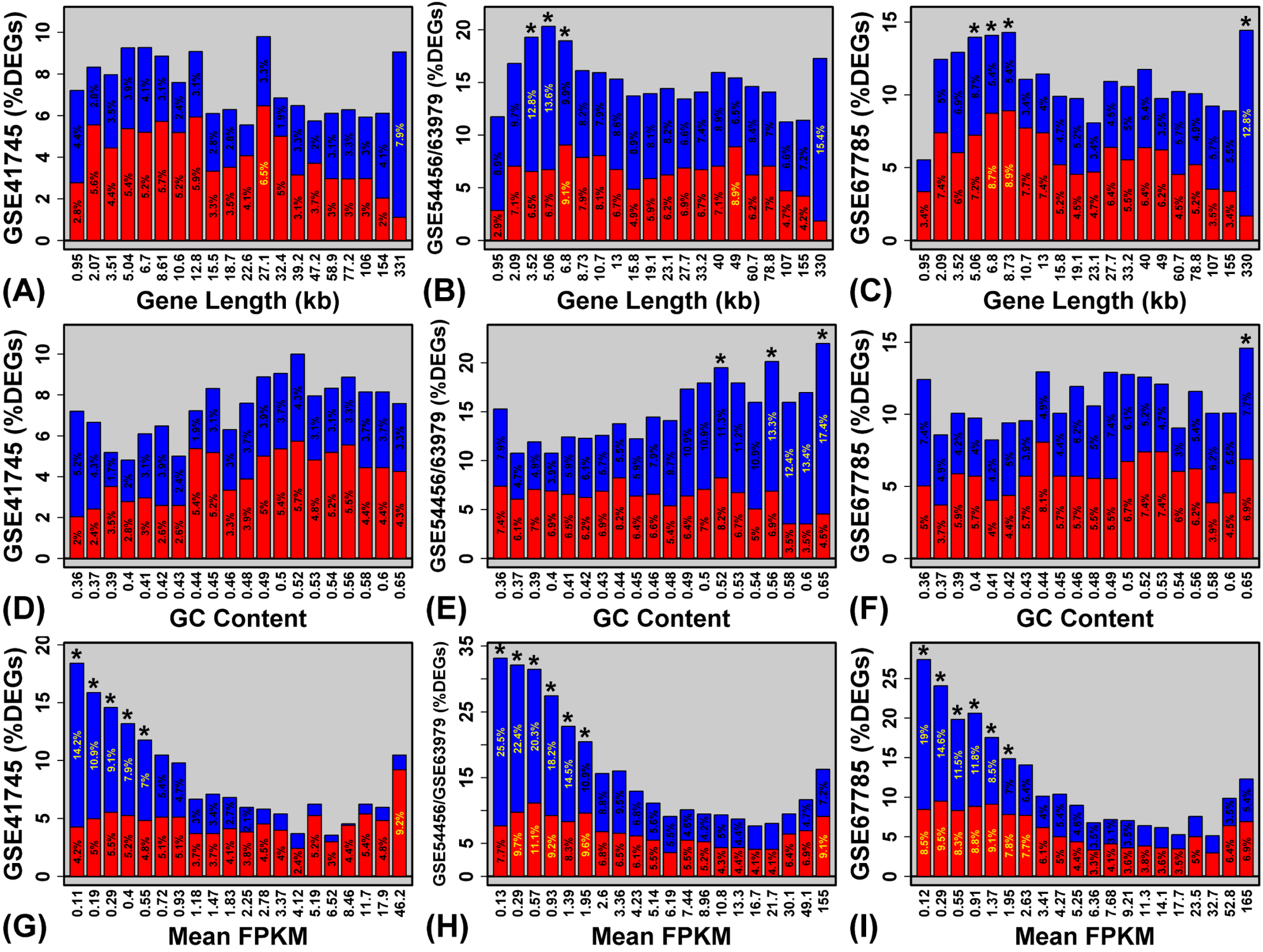

GSE41745 (%DEGs)
GSE54456/63979 (%DEGs)
GSE67785 (%DEGs)
Gene Length (kb)
GC Content
Mean FPKM
(A)
(B)
(C)
(D)
(E)
(F)
(G)
(H)
(I)

Figure S6

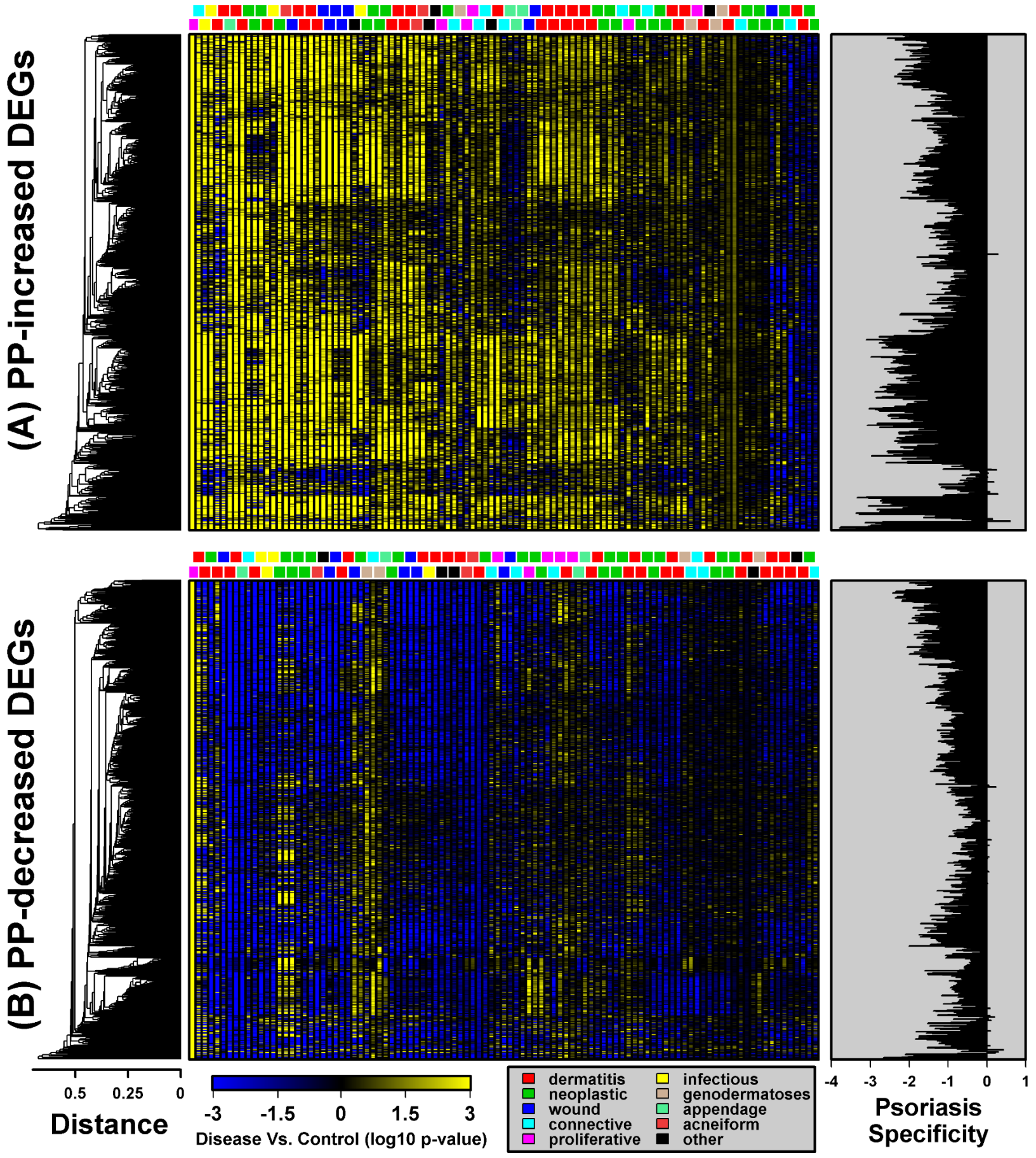

(A) PP-increased DEGs
(B) PP-decreased DEGs
0.5
0.25
0
Distance
-3
-1.5
0
1.5
3
Disease Vs. Control (log10 p-value)
dermatitis
neoplastic
wound
connective
proliferative
infectious
genodermatoses
appendage
acneiform
other
-4
-3
-2
-1
0
1
Psoriasis Specificity

Figure S7

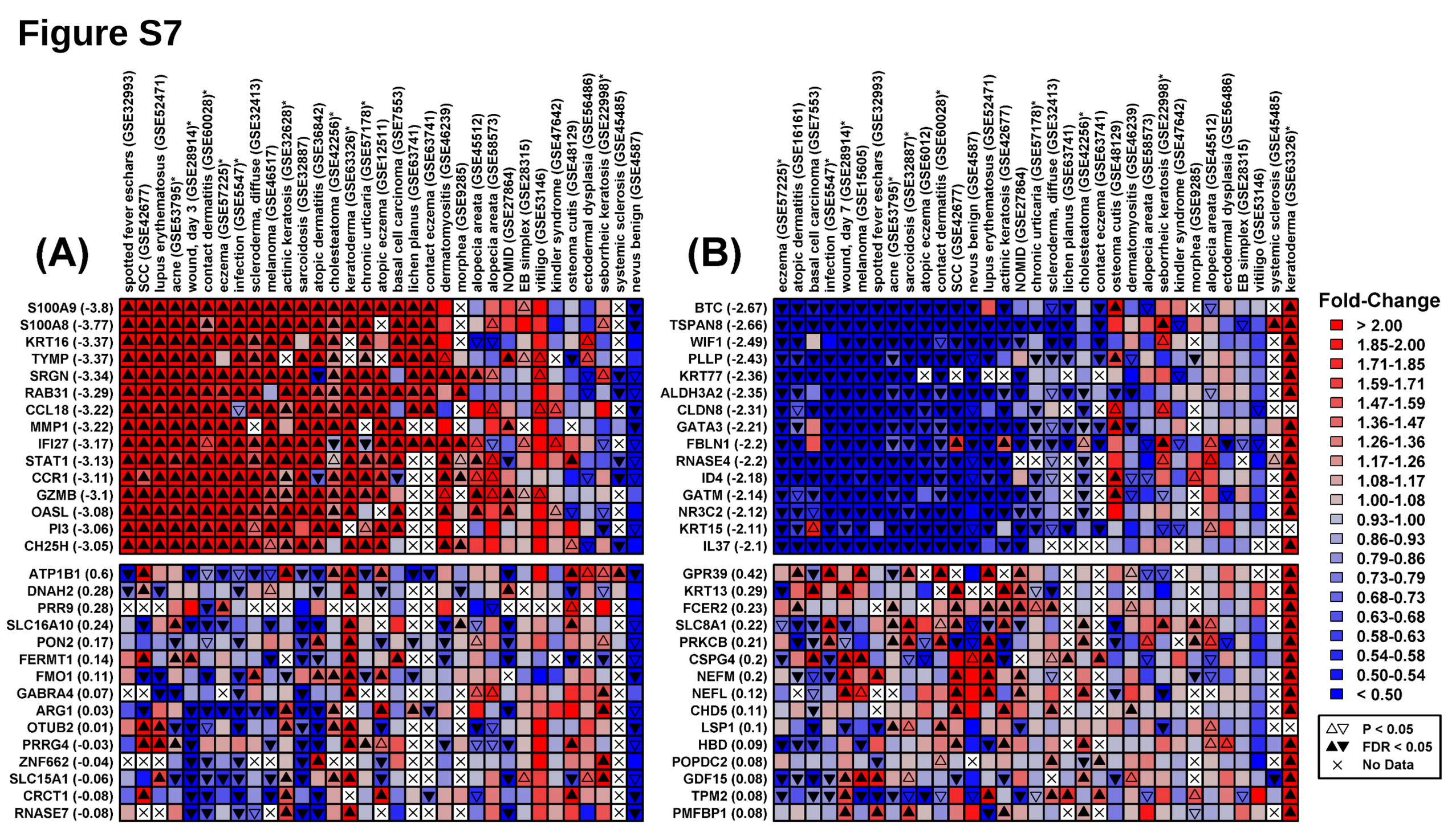
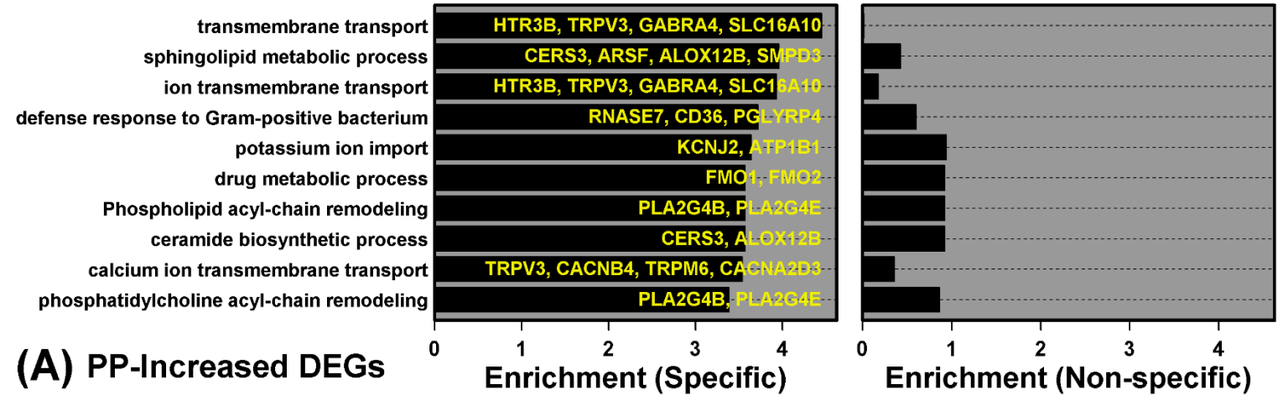
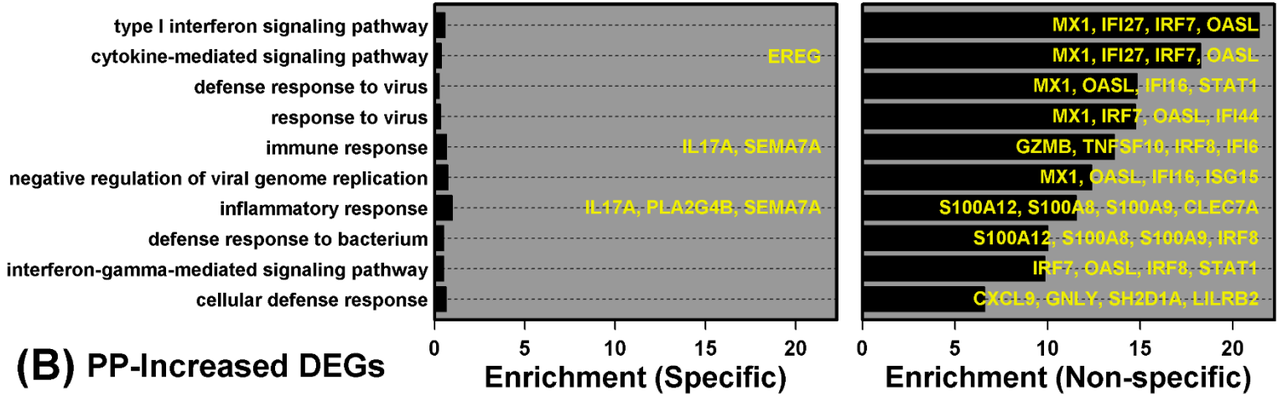

(A)
spotted fever eschars (GSE32993)
SCC (GSE42677)
lupus erythematosus (GSE52471)
acne (GSE53795)*
wound, day 3 (GSE28914)*
contact dermatitis (GSE60028)*
eczema (GSE57225)*
infection (GSE5547)*
scleroderma, diffuse (GSE32413)
melanoma (GSE46517)
actinic keratosis (GSE32628)*
sarcoidosis (GSE32887)
atopic dermatitis (GSE36842)
cholesteatoma (GSE42256)*
keratoderma (GSE63326)*
chronic urticaria (GSE57178)*
atopic eczema (GSE12511)
basal cell carcinoma (GSE7553)
lichen planus (GSE63741)
contact eczema (GSE63741)
dermatomyositis (GSE46239)
morphea (GSE9285)
alopecia areata (GSE45512)
alopecia areata (GSE58573)
NOMID (GSE27864)
EB simplex (GSE28315)
vitiligo (GSE53146)
kindler syndrome (GSE47642)
osteoma cutis (GSE48129)
ectodermal dysplasia (GSE56486)
seborrheic keratosis (GSE22998)*
systemic sclerosis (GSE45485)
nevus benign (GSE4587)
S100A9 (-3.8)
S100A8 (-3.77)
KRT16 (-3.37)
TYMP (-3.37)
SRGN (-3.34)
RAB31 (-3.29)
CCL18 (-3.22)
MMP1 (-3.22)
IFI27 (-3.17)
STAT1 (-3.13)
CCR1 (-3.11)
GZMB (-3.1)
OASL (-3.08)
PI3 (-3.06)
CH25H (-3.05)
ATP1B1 (0.6)
DNAH2 (0.28)
PRR9 (0.28)
SLC16A10 (0.24)
PON2 (0.17)
FERMT1 (0.14)
FMO1 (0.11)
GABRA4 (0.07)
ARG1 (0.03)
OTUB2 (0.01)
PRRG4 (-0.03)
ZNF662 (-0.04)
SLC15A1 (-0.06)
CRCT1 (-0.08)
RNASE7 (-0.08)
(B)
eczema (GSE57225)*
atopic dermatitis (GSE16161)
basal cell carcinoma (GSE7553)
infection (GSE5547)*
wound, day 7 (GSE28914)*
melanoma (GSE15605)
spotted fever eschars (GSE32993)
acne (GSE53795)*
sarcoidosis (GSE32887)*
atopic eczema (GSE6012)
contact dermatitis (GSE60028)*
SCC (GSE42677)
nevus benign (GSE4587)
lupus erythematosus (GSE52471)
actinic keratosis (GSE42677)
NOMID (GSE27864)
chronic urticaria (GSE57178)*
scleroderma, diffuse (GSE32413)
lichen planus (GSE63741)
cholesteatoma (GSE42256)*
contact eczema (GSE63741)
osteoma cutis (GSE48129)
dermatomyositis (GSE46239)
alopecia areata (GSE58573)
seborrheic keratosis (GSE22998)*
kindler syndrome (GSE47642)
morphea (GSE9285)
alopecia areata (GSE45512)
ectodermal dysplasia (GSE56486)
EB simplex (GSE28315)
vitiligo (GSE53146)
systemic sclerosis (GSE45485)
keratoderma (GSE63326)*
BTC (-2.67)
TSPAN8 (-2.66)
WIF1 (-2.49)
PLLP (-2.43)
KRT77 (-2.36)
ALDH3A2 (-2.35)
CLDN8 (-2.31)
GATA3 (-2.21)
FBLN1 (-2.2)
RNASE4 (-2.2)
ID4 (-2.18)
GATM (-2.14)
NR3C2 (-2.12)
KRT15 (-2.11)
IL37 (-2.1)
GPR39 (0.42)
KRT13 (0.29)
FCER2 (0.23)
SLC8A1 (0.22)
PRKCB (0.21)
CSPG4 (0.2)
NEFM (0.2)
NEFL (0.12)
CHD5 (0.11)
LSP1 (0.1)
HBD (0.09)
POPDC2 (0.08)
GDF15 (0.08)
TPM2 (0.08)
PMFBP1 (0.08)
Fold-Change
> 2.00
1.85-2.00
1.71-1.85
1.59-1.71
1.47-1.59
1.36-1.47
1.26-1.36
1.17-1.26
1.08-1.17
1.00-1.08
0.93-1.00
0.86-0.93
0.79-0.86
0.73-0.79
0.68-0.73
0.63-0.68
0.58-0.63
0.54-0.58
0.50-0.54
< 0.50
P < 0.05
FDR < 0.05
No Data

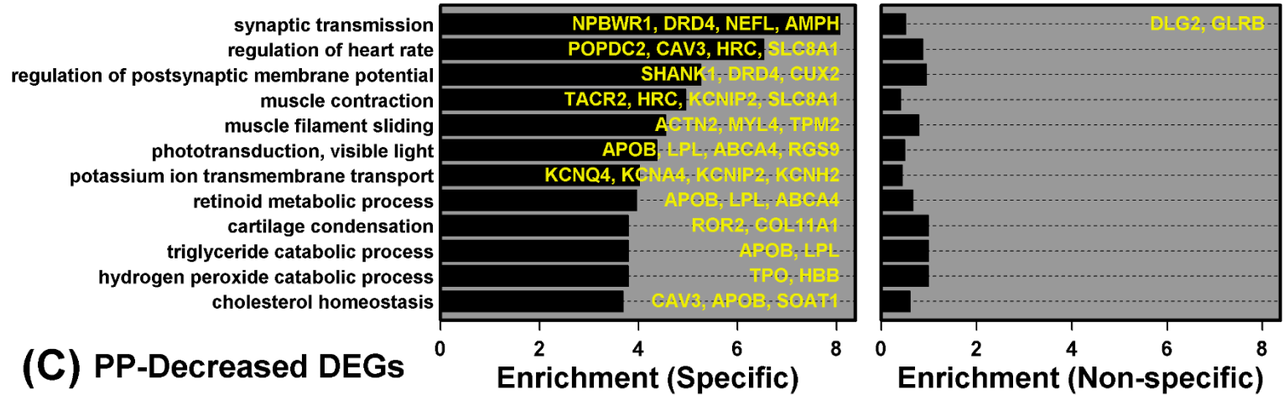
Figure S8



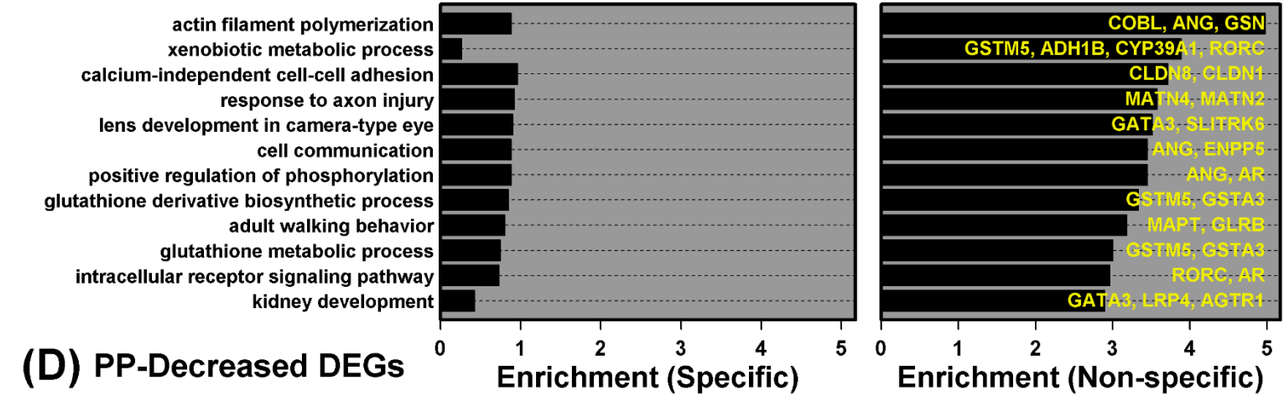
(A) PP-Increased DEGs



(B) PP-Increased DEGs

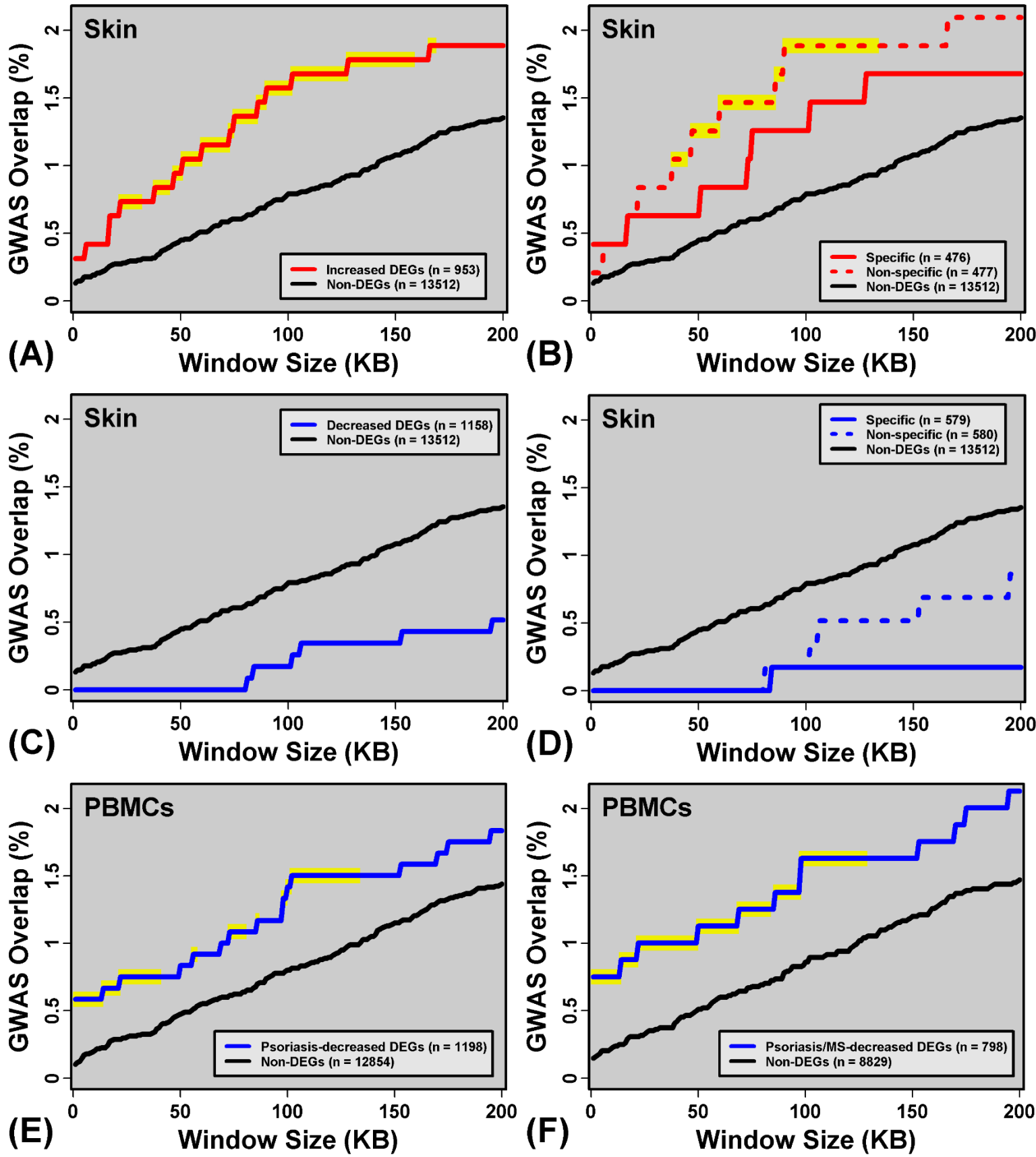


(C) PP-Decreased DEGs



(D) PP-Decreased DEGs

Figure S9


Skin
GWAS Overlap (%)
Window Size (KB)
Increased DEGs (n = 953)
Non-DEGs (n = 13512)
(A)
Skin
Specific (n = 476)
Non-specific (n = 477)
Non-DEGs (n = 13512)
(B)
Skin
Decreased DEGs (n = 1158)
Non-DEGs (n = 13512)
(C)
Skin
Specific (n = 579)
Non-specific (n = 580)
Non-DEGs (n = 13512)
(D)
PBMCs
Psoriasis-decreased DEGs (n = 1198)
Non-DEGs (n = 12854)
(E)
PBMCs
Psoriasis/MS-decreased DEGs (n = 798)
Non-DEGs (n = 8829)
(F)