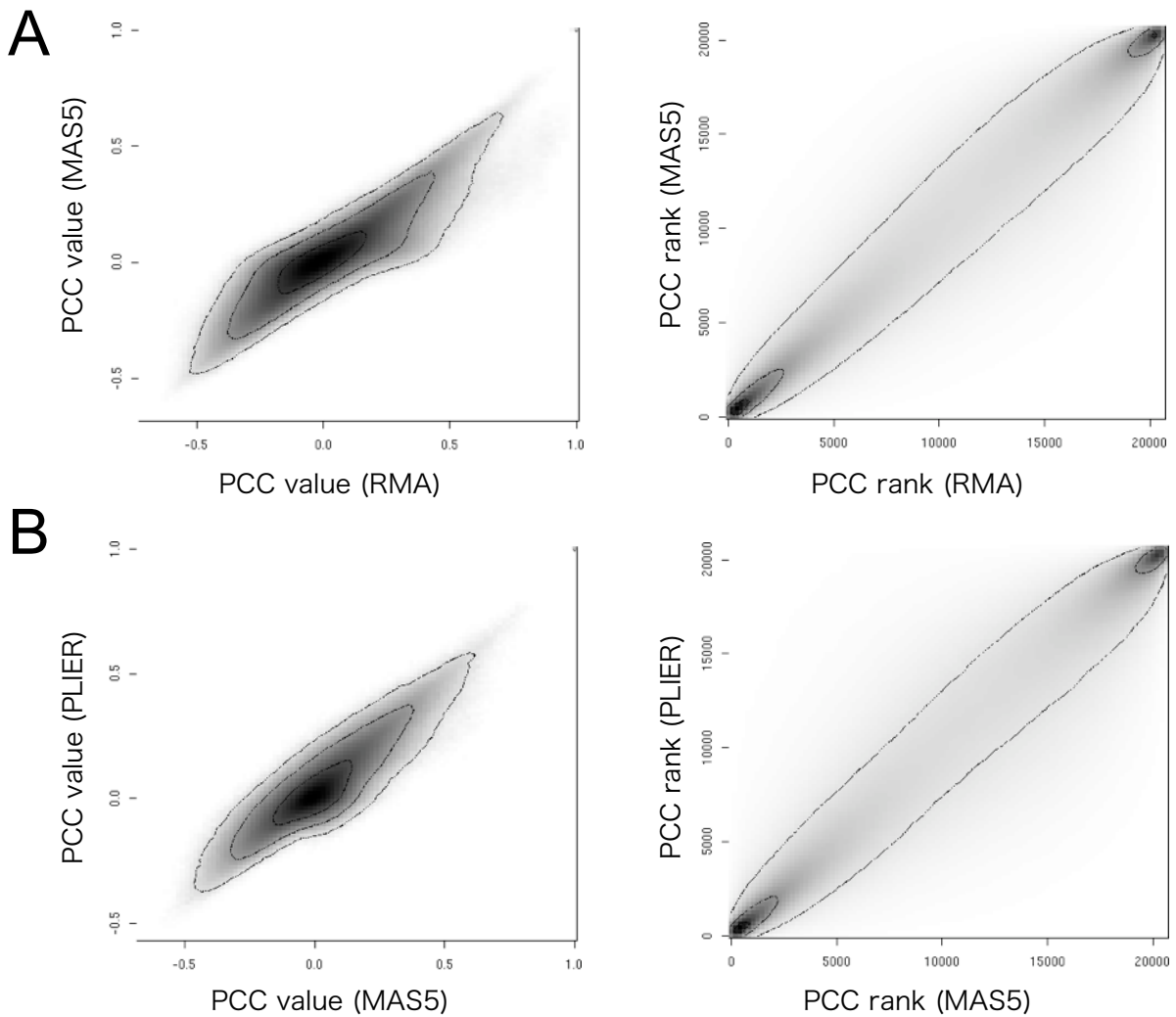


### Supplementary Figure S1

ROC curves for prediction of gene function using GO terms. Because the shapes of the curves were standard convex-upward, we used the area under the ROC curve (i.e., AUC) to compare various coexpression data in this manuscript. BP, biological process; CC, cellular component; MF, molecular function.

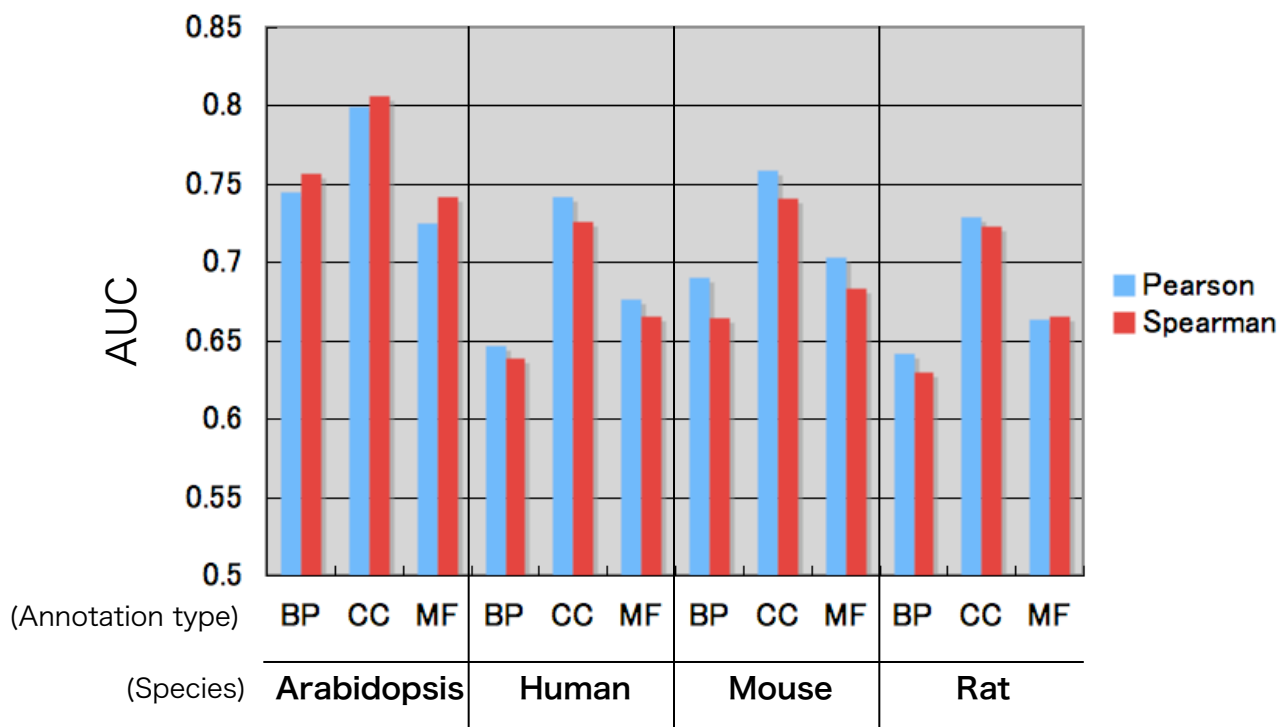


### Supplementary Figure S2

Inconsistency of the PCC values and consistency of the PCC ranks for differently constructed expression data.

Degree of coexpression is compared using PCC value (left panels) or PCC rank (right panels). (A) RMA vs. MAS5, (B) MAS5 vs. PLIER.

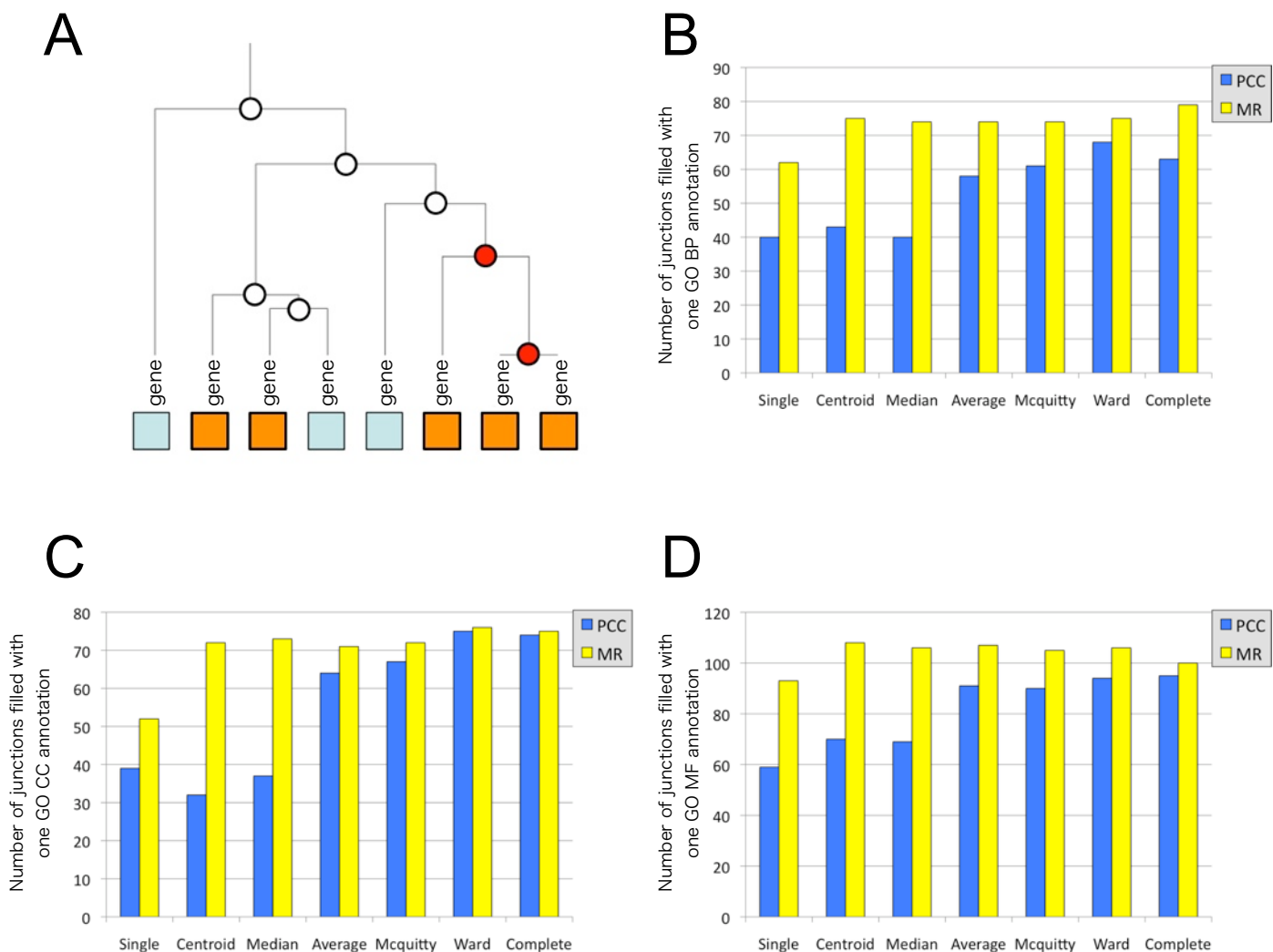
Distributions are represented as density plots in logarithmic scale.



### Supplementary Figure S3

Comparison of PCC and Spearman Correlation Coefficient as a coexpression measure.

AUCs of the prediction of GO terms are shown. GO annotations: BP, biological process; CC, cellular component; MF, molecular function.



### Supplementary Figure S4

Comparison of PCC and MR for hierarchical gene clustering.

(A) Schematic figure of the method to assess the goodness of the obtained clusters. The junction nodes filled with just one GO annotation were counted. The orange and cyan boxes indicate GO annotations. The two red junction nodes are filled with only the orange GO annotation, while the other junction has genes with two GO annotations. Therefore, in this example, the number of junctions with the same GO annotation is two. Noted that we only used genes with one GO annotation for the clustering. (B) Assessment of the gene clustering in BP category. Y-axis is the number of the junctions filled with a single BP term. Blue and yellow bars indicate the performance for PCC values and MR with the seven different linkage rules. (C) Same as B but for CC category. (D) Same as B but for MF category.

## Supplementary Data (A): Data source detail for Arabidopsis.

| TAIR ID       | Sample background     | Number of slides | Sample description   |
|---------------|-----------------------|------------------|--|
| TAIR-ME00319  |                       | 237              | AtGenExpress: Expression Atlas of Arabidopsis Development  |
| TAIR-ME00346  |                       | 6                | AtGenExpress: Expression profiling of early germinating seeds  |
| TAIR-ME00341  |                       | 12               | AtGenExpress: Response to Botrytis cinerea infection   |
| TAIR-ME00353  |                       | 32               | AtGenExpress: Pseudomonas half leaf injection  |
| TAIR-ME00331  |                       | 48               | AtGenExpress: Response to virulent, avirulent, typeIII-secretion system deficient and nonhost bacteria |
| TAIR-ME00332  |                       | 42               | AtGenExpress: Response to bacterial-(LPS, HrpZ, Flg22) and oomycete-(NPP1) derived elicitors           |
| TAIR-ME00342  |                       | 18               | AtGenExpress: Response to Phytophthora infestans   |
| TAIR-ME00354  |                       | 48               | AtGenExpress: Response to Erysiphe orontii infection   |
| TAIR-ME00325  | shoot                 | 26               | AtGenExpress: Cold stress time course  |
|               | root                  | 26               |  |
| TAIR-ME00338  | shoot                 | 30               | AtGenExpress: Drought stress time course   |
|               | root                  | 30               |  |
| TAIR-ME00326  | shoot                 | 26               | AtGenExpress: Genotoxic stress time course   |
|               | root                  | 26               |  |
| TAIR-ME00339  | shoot                 | 34               | AtGenExpress: Heat stress time course  |
|               | root                  | 34               |  |
|               | cell suspension cultu | 26               |  |
| TAIR-ME00327  | shoot                 | 26               | AtGenExpress: Osmotic stress time course   |
|               | root                  | 26               |  |
| TAIR-ME00340  | shoot                 | 26               | AtGenExpress: Oxidative stress time course   |
|               | root                  | 26               |  |
| TAIR-ME00328  | shoot                 | 26               | AtGenExpress: Salt stress time course  |
|               | root                  | 26               |  |
| TAIR-ME00329  | shoot                 | 30               | AtGenExpress: UV-B stress time course  |
|               | root                  | 30               |  |
| TAIR-ME00330  | shoot                 | 30               | AtGenExpress: Wounding stress time course  |
|               | root                  | 30               |  |
| TAIR-ME00345  |                       | 48               | AtGenExpress: Light treatments   |
| TAIR-ME00348  |                       | 4                | AtGenExpress: Different temperature treatment of seeds   |
| TAIR-ME00347  |                       | 22               | AtGenExpress: Response to sulfate limitation   |
| TAIR-ME00334  |                       | 12               | AtGenExpress: ACC time course in wildtype seedlings  |
| TAIR-ME00344  |                       | 12               | AtGenExpress: Zeatin time course in wildtype seedlings   |
| TAIR-ME00337  |                       | 12               | AtGenExpress: Methyl Jasmonate time course in wildtype seedlings                                       |
| TAIR-ME00336  |                       | 12               | AtGenExpress: IAA time course in wildtype seedlings  |
| TAIR-ME00333  |                       | 12               | AtGenExpress: ABA time course in wildtype seedlings  |
| TAIR-ME00343  | wild type             | 12               | AtGenExpress: GA3 time course in wildtype and ga1-5 mutant seedlings                                   |
|               | mutant                | 12               |  |
| TAIR-ME00335  | wild type             | 12               | AtGenExpress: Brassinolide time course in wildtype and det2-1 mutant seedlings                         |
|               | mutant                | 12               |  |
| TAIR-ME00352  |                       | 26               | AtGenExpress: Effect of Brassinosteroids on seedlings  |
| TAIR-ME00356  |                       | 12               | AtGenExpress: Cytokinin treatment of seedlings   |
| TAIR-ME00351  |                       | 8                | AtGenExpress: Effect of ABA during seed imbibition   |
| TAIR-ME00350  |                       | 12               | AtGenExpress: Basic hormone treatment of seeds   |
| TAIR-ME00357  |                       | 20               | AtGenExpress: Effect of gibberellic acid inhibitors on seedlings                                       |
| TAIR-ME00358  |                       | 10               | AtGenExpress: Effect of auxin inhibitors on seedlings  |
| TAIR-ME00359  |                       | 18               | AtGenExpress: Effect of brassinosteroid inhibitors on seedlings  |
| TAIR-ME00360  |                       | 6                | AtGenExpress: Effect of ethylene inhibitors on seedlings   |
| TAIR-ME00361  |                       | 4                | AtGenExpress: Effect of cycloheximide on seedlings   |
| TAIR-ME00362  |                       | 4                | AtGenExpress: Effect of proteasome inhibitor MG13 on seedlings   |
| TAIR-ME00363  |                       | 12               | AtGenExpress: Effect of photosystem inhibitor PNO8 on seedlings  |
| TAIR-ME00364  |                       | 8                | AtGenExpress: Effect of ibuprofen, salicylic acid and daminozide on seedlings                          |
| TAIR-ME00365  |                       | 10               | AtGenExpress: Genome-wide cell cycle studies   |
| TAIR-ME00367  |                       | 4                | AtGenExpress: Growth of suspension-cultured cells  |
| TAIR-ME00366x |                       | 7                | AtGenExpress: Cell cycle re-entry in suspension-cultured cells   |
| TAIR-ME00349  |                       | 8                | AtGenExpress: Comparison of plant hormone-related mutants  |
| TAIR-ME00355  |                       | 6                | AtGenExpress: ARR21C overexpression  |
| TAIR-ME00374  |                       | 30               | AtGenExpress: Ecotypes   |
| TAIR-ME00375  |                       | 24               | AtGenExpress: Ecotypes Singletons  |

These data are downloaded from TAIR web site. <http://www.arabidopsis.org/info/expression/ATGenExpress.jsp>

## Supplementary Data (B): Data source detail for Human

| GSE ID  | Subgroup for normalization | Number of slides | Experimental type                              | Sample description  |
|---------|----------------------------|------------------|--|---|
| GSE2109 | batch01                    | 60               | cancer portraits                               | Expression Project for Oncology (expO)  |
|         | batch02                    | 163              |  |   |
|         | batch03                    | 164              |  |   |
|         | batch04                    | 163              |  |   |
|         | batch05                    | 160              |  |   |
|         | batch06                    | 160              |  |   |
|         | batch07                    | 156              |  |   |
|         | batch08                    | 160              |  |   |
|         | batch09                    | 240              |  |   |
|         | batch10                    | 180              |  |   |
|         | batch11                    | 180              |  |   |
|         | batch12                    | 125              |  |   |
|         | batch13                    | 62               |  |   |
| GSE2125 |                            | 45               | parallel sample                                | isolated alveolar macrophages   |
| GSE2328 |                            | 9                | other  | Application of genome-wide expression analysis to human health & disease  |
| GSE2435 |                            | 6                | ordered  | Autophagy promotes MHC class II presentation of peptides from intracellular source proteins                             |
| GSE2555 |                            | 2                | parallel sample                                | HCaRG-9 vs NEO-1  |
| GSE2600 |                            | 6                | ordered  | Anaplasma phagocytophilum infected NB4 cells  |
| GSE2634 |                            | 17               | parallel sample                                | Comparison of human and non-human primate gene expression profiles  |
| GSE2677 |                            | 39               | other  | Prednisolone treated childhood ALL samples  |
| GSE2816 |                            | 6                | repeat   | cMyb and vMyb in human monocytes  |
| GSE2817 |                            | 30               | wavelet, glioma, unsupervised                  | Wavelet modelling of microarray data provides chromosomal pattern of expression which predicts survival in gliomas      |
| GSE2842 |                            | 45               | treatment response                             | Additional systems to Prednisolone treated childhood ALL samples  |
| GSE2864 |                            | 6                | time-course                                    | tour-affy-human-131547  |
| GSE3077 |                            | 12               | Dilution series, platform comparison, Affyr    | Dilution series comparison of Affymetrix and Illumina Expression Platforms  |
| GSE3202 |                            | 23               | drug treatment                                 | MK886 treatment of H720 non-small cell lung cancer cell line  |
| GSE3284 |                            | 18               | Gene expression profiling of human blood       | A network-based analysis of systemic inflammation in humans   |
| GSE3325 |                            | 19               | prostate cancer progression                    | Integrative Genomic and Proteomic Analysis of Prostate Cancer Reveals Signatures of Metastatic Progression              |
| GSE3526 |                            | 353              | normal human tissue comparison                 | Comparison of gene expression profiles across the normal human body   |
| GSE3678 |                            | 14               | gene expression analysis                       | PTC versus paired normal thyroid tissue   |
| GSE3744 |                            | 47               | Type   | Human breast tumor expression   |
| GSE4030 |                            | 8                | time-course                                    | bunge-affy-arabi-162779   |
| GSE4036 |                            | 28               | other  | perro-affy-human-186940   |
| GSE4064 |                            | 4                | tissue type and state comparison               | Gene expression of growing antler in Cervus elaphus   |
| GSE4086 |                            | 4                | Hypoxia responsive, case control               | Hypoxia responsive genes in human Burkitt's lymphoma cell line, P493-6.   |
| GSE4182 |                            | 9                | Prenatal gene expression signature             | Genetic background of the polygenic neural tube defect  |
| GSE4217 |                            | 12               | other  | Spheroid Formation and Recovery of Human Foreskin Fibroblasts at Ambient Temperature                                    |
| GSE4218 |                            | 12               | other  | Spheroid Formation and Recovery of Human T98G Glioma Cells at Ambient Temperature                                       |
| GSE4219 |                            | 24               | SuperSeries                                    | Spheroid Formation and Recovery of Human Foreskin Fibroblasts and T98G Glioma Cells at Ambient Temperature              |
| GSE4237 |                            | 10               | other  | Hussaini-2R01NS035122-06A1  |
| GSE4250 |                            | 4                | tissue specific expression profile             | Determination of expression level of gingiva expressed genes  |
| GSE4488 |                            | 7                | affected and obligatory carriers vs control    | Expression data from whole blood  |
| GSE4567 |                            | 8                | particle treatment                             | Endothelial cell culture with Chapel Hill Ultrafine particle  |
| GSE4600 |                            | 11               | Neuronal Differentiation, Targets of MeCP      | Identifying targets of MeCP2 during neuronal maturation differentiation   |
| GSE4737 |                            | 2                | parallel sample                                | HCaRG vs NEO  |
| GSE4757 |                            | 20               | other  | Rogers-3U24NS043571-01S1  |
| GSE4773 |                            | 20               | time-course                                    | Greene-5P01NS017771-220003  |
| GSE4780 |                            | 57               | other  | Scheck-3U24NS043571-01S1-2  |
| GSE4824 |                            | 6                | cell type comparison                           | Analysis of lung cancer cell lines  |
| GSE4883 |                            | 9                | time-course, response to treatment             | Simvastatin has an anti-inflammatory effect on macrophages via upregulation of Kruppel-like factor-2                    |
| GSE4888 |                            | 27               | Gene expression arrays human endometri         | Molecular phenotyping of human endometrium  |
| GSE4984 |                            | 12               | dendritic cell maturation                      | Monocyte Derived Dendritic Cell Maturation  |
| GSE5109 |                            | 6                | Time-Series                                    | Gastric Bypass Human Obese Muscle   |
| GSE5110 |                            | 10               | stimulus or stress design                      | 48h Immobilization in human   |
| GSE5145 |                            | 6                | comparison                                     | Genes regulated by vitamin D in bronchial smooth muscle cells   |
| GSE5264 |                            | 30               | time course                                    | Human bronchial epithelial cell differentiation time course   |
| GSE5281 |                            | 74               | other  | steph-affy-human-433773   |
| GSE5350 |                            | 140              | Cross-platform comparison                      | MicroArray Quality Control (MAQC) Project   |
| GSE5372 |                            | 22               | response to airway injury                      | airway epithelium, large airways, pre and post-mechanical injury  |
| GSE5486 |                            | 8                | microarray, gene expression, Affymetrix, d     | Using GIN2 to identify novel mutations in candidate tumor suppressor genes in colon cancer cells                        |
| GSE5563 |                            | 19               | disease state analysis                         | Gene expression profile of VIN lesions in comparison to controls  |
| GSE5675 |                            | 41               | Human, WHO grade I, BRain tumor, Pilocyt       | Pilocytic astrocytoma   |
| GSE5679 |                            | 11               | ligand response                                | Comparative gene expression profile of PPARg and RARa ligand treated human dendritic cells                              |
| GSE5741 |                            | 9                | treatment with non-treatment control           | Expression data from ARPE-19 cells treated with LDL or ox-LDL   |
| GSE5764 |                            | 30               | cell type comparison                           | Analysis of microdissected invasive lobular and ductal breast carcinomas in relation to normal ductal and lobular cells |
| GSE5787 |                            | 33               | disease state comparison                       | Survey of Intra- and Inter-Tumour Heterogeneity of Gene Expression in Cervical Cancer                                   |
| GSE5809 |                            | 14               | Gene expression arrays in human stromal        | Decidual stromal cell response to paracrine signals from the trophoblast  |
| GSE5816 |                            | 29               | Cell line comparison                           | A Genome-wide Screen for Hypermethylated Genes in Lung Cancer   |
| GSE5823 |                            | 20               | siRNA analysis                                 | Effects of RNA interference-mediated c-MYC depletion on gene expression profiles in human cancer cell lines             |
| GSE5968 |                            | 12               | ERRalpha, PGC-1alpha, nuclear receptor,        | Identification of ERRalpha dependent PGC-1alpha target genes  |
| GSE5993 |                            | 6                | differential expression profiling from trans   | mRNA expression profiling of p63-depleted versus wildtype ME180 cells   |
| GSE6004 |                            | 18               | Genetic modification                           | Gene Expression and Functional Evidence of Epithelial-to-Mesenchymal Transition in Papillary Thyroid Cancer Invasion    |
| GSE6013 |                            | 27               | time course, gene expression, cell line, ast   | Gene expression in asbestos exposed lung cells  |
| GSE6021 |                            | 4                | RNA immunoprecipitation                        | Immunoprecipitation of PTB65 associated mRNAs   |
| GSE6022 |                            | 6                | RNA immunoprecipitation                        | Immunoprecipitation of U2AF65 associated mRNAs  |
| GSE6075 |                            | 3                | oncogenesis.                                   | E6 in HOK/Bmi-1 cells   |
| GSE6090 |                            | 6                | Activation state, signalling, Toll-like Recept | DC-SIGN initiates an immature dendritic cell phenotype triggering Rho activation that is utilised by HIV-1              |
| GSE6092 |                            | 8                | Cell response to inflammatory stimuli          | IFN-gamma alters the response of Borrelia burgdorferi-activated endothelium to favor chronic inflammation               |
| GSE6140 |                            | 6                | platform comparison                            | Cross platform microarray analysis for robust identification of differentially expressed genes                          |
| GSE6207 |                            | 14               | time course                                    | miR-124 transfection time course  |
| GSE6233 |                            | 2                | Gene expression after RNAi                     | Expression data in ezrin-knockdown ec109 cells(PSE1) and control(PS)  |
| GSE6241 |                            | 8                | time course                                    | The effects of Serum Amyloid A on gene expression profile in HUVECs   |
| GSE6257 |                            | 6                | TNFalpha, Lymphatic endothelium, inducti       | An inflammation-induced mechanism for leukocyte transmigration of lymphatic vessel endothelium.                         |
| GSE6269 |                            | 22               | expression analysis                            | Gene expression patterns in blood leukocytes discriminate patients with acute infections                                |
| GSE6276 |                            | 6                | dose response                                  | rogae-affy-human-323460   |
| GSE6281 |                            | 34               | time course, allergic response                 | Gene expression time-course in the human skin during elicitation of allergic contact dermatitis                         |
| GSE6283 |                            | 21               | Trisomy, Down syndrome, Patau syndrom          | Specific transcriptional networks in human fetus with autosomal trisomies   |
| GSE6338 |                            | 60               | Molecular pathogenesis, molecular classifi     | Gene expression analysis of Peripheral T-cell Lymphoma/Unspecified  |
| GSE6357 |                            | 6                | Renal cell carcinoma, CD8+ T cells, micro:     | Activation of human CD8+ T cells in renal cell carcinoma  |

|         |         |     |  |  |
|---------|---------|-----|--|--|
| GSE6364 |         | 37  | disease state analysis                       | Gene Profiling of Endometrium Reveals Progesterone Resistance and Candidate Genetic Loci in Women with Endometriosis     |
| GSE6369 |         | 2   | disease-state analysis                       | Prostate Adenocarcinoma Progression  |
| GSE6400 |         | 12  | responses to treatments with anti-cancer     | Cultured A549 lung cancer cells treated with actinomycin D and saphyrin PCI-2050   |
| GSE6451 |         | 6   | overexpression of different antioxidant res  | A genomic screen for activators of the antioxidant response element  |
| GSE6495 |         | 6   | comparative genomic hybridization            | NOTCH signaling in T-cell acute lymphoblastic leukemia cell lines  |
| GSE6519 |         | 12  | dietary dose response                        | Microarray Analysis of Baboon neonates consuming long-chain polyunsaturated fatty acid formula                           |
| GSE6521 |         | 36  | time course                                  | MCF7 inhibitor   |
| GSE6532 |         | 87  | disease state analysis                       | Definition of clinically distinct molecular subtypes in estrogen receptor positive breast carcinomas using genomic grade |
| GSE6561 |         | 4   | karyotypically abnormal hESC vs karyotyp     | Expression data from hsr containing human embryonic stem cell line, H14, and parent H14 human embryonic stem cell line.  |
| GSE6573 |         | 6   | disease comparison, disease state            | Dysregulation of the circulating and tissue-based renin-angiotensin system in preeclampsia                               |
| GSE6575 |         | 56  | autism analysis                              | Gene expression in blood of children with autism spectrum disorder   |
| GSE6653 |         | 8   | Time-course                                  | Gene expression analysis of IQSE cells treated with TGFb1, a time course study   |
| GSE6743 |         | 6   | Human T cell, vitamin D and A, chemokine     | 1,25 (OH)2 vitamin D3 induces expression of CCR10 and other genes  |
| GSE6791 |         | 81  | Gene expression ptofiles of primary cance    | Gene Expression Profiles of HPV-Positive and -Negative Head/Neck and Cervical Cancers                                    |
| GSE6816 |         | 2   | Melanotransferrin, hyperepression cell lin   | Hyperexpression of Human Melanotransferrin on SK-N-MC Cell Line  |
| GSE6817 |         | 3   | Melanotransferrin, siRNA, comparative ger    | Downregulation of Human Melanotransferrin on SK-Mel-28 Cell Line   |
| GSE6819 |         | 6   | siRNA treatment study                        | Identification of genes that are linked with optineurin expression   |
| GSE6833 |         | 5   | SuperSeries                                  | Hyperexpression and Downregulation of Melanotransferrin on Various Cell Lines  |
| GSE6863 |         | 6   | stress response, cell differentiation        | Immature dendritic cells under hypoxic condition   |
| GSE6872 |         | 21  | disease state analysis                       | Spermatozoal RNA Profiles (U133 Plus 2.0 Array)  |
| GSE6890 |         | 5   | cell line comparison                         | Transcriptome maps of six different human cell lines   |
| GSE6960 |         | 12  | Dose response                                | Synthesis and Anticancer Properties of Water-Soluble Zinc Ionophores 1   |
| GSE6962 |         | 12  | Dose response                                | Synthesis and Anticancer Properties of Water-Soluble Zinc Ionophores 2   |
| GSE6965 |         | 4   | dendritic cell, A. fumigatus germ tubes, inf | Gene expression profil of human dendritic cells after infection with A. fumigatus  |
| GSE6969 |         | 21  | SuperSeries                                  | Spermatozoal RNA Profiles from Normally Fertile and Heterogeneously Teratozoospermic Subjects                            |
| GSE6972 |         | 24  | SuperSeries                                  | Synthesis and Anticancer Properties of Water-Soluble Zinc Ionophore: Cell Culture and Xenograft Model                    |
| GSE7036 |         | 6   | Bipolar Disorder, Monozygotic Twins, Disc    | Expression profiling in monozygotic twins discordant for bipolar disorder  |
| GSE7127 |         | 62  | Expression profiling by array                | 63 Melanoma cell lines   |
| GSE7153 |         | 17  | Affymetrix Hu133,Plus microarrays            | Transcriptional targets of p16INK4A in melanoma  |
| GSE7181 |         | 6   | Cancer stem cell, CD133, glioblastoma        | CD133+ and CD133- CSC lines in primary human GBM   |
| GSE7213 |         | 4   | comparative expression profiling             | Comparative analysis of cellular mRNA incorporation into MLV and HIV1 virus-like particles                               |
| GSE7216 |         | 25  | cytokine response                            | Cytokine treated normal human epidermal keratinocytes  |
| GSE7234 |         | 4   | different karyotype                          | Expression data from different karyotypic human embryonic stem cells   |
| GSE7247 |         | 10  | Loading methodology comparison               | Dendritic Cells Compare the Similarity of Endogenous and Exogenous Antigens  |
| GSE7305 |         | 20  | Human tissue transcriptional profiling       | Human endometriosis vs normal endometrium study - transcriptional profiling  |
| GSE7307 | devide1 | 339 | Human body index of gene expression          | Human body index - transcriptional profiling   |
|         | devide2 | 338 |  |  |
| GSE7364 |         | 8   | genomic DNA, CGH, Copy Number Variati        | Gene-resolution analysis of DNA copy number variation using oligonucleotide expression microarrays                       |
| GSE7382 |         | 6   | Inducible cell lines                         | Expression profiling of MCF-7 cells with inducible LMO4 and DN-Clim expression   |
| GSE7392 |         | 30  | time course, comparative expression          | Molecular Evidence of Injury and Inflammation in Normal and Fibrotic Renal Allografts One Year Post-Transplant           |
| GSE7508 |         | 6   | effects of gene knockdown                    | Identification of LEDGF-responsive genes in Jurkat cells   |
| GSE7509 |         | 26  | response to antibody                         | Gene expression changes in anti-FcgRIIb treated DCs and monocytes  |
| GSE7554 |         | 8   | comparison, experiential versus control      | Identification of Cellular Genes Targeted by KSHV-Encoded MicroRNAs  |
| GSE7578 |         | 15  | shRNA knock-down                             | shRNA-mediated knock down of Bmi-1 and Mel-18 in medulloblastoma cells   |
| GSE7708 |         | 14  | Gene expression changes in cultured LNC      | Suppression of androgen receptor mediated gene expression by a sequence-specific DNA binding polyamide                   |
| GSE7745 |         | 6   | CHIP-CHIP and expression data                | Mapping of HNF4b binding sites, acetylation of histone H3 and expression in Caco2 cells                                  |
| GSE7835 |         | 15  | Gene expression changes in cultured U25      | Modulating hypoxia-inducible transcription by disrupting the HIF-1-DNA interface   |
| GSE7845 |         | 2   | Comparative genomic hybridization            | Proinflammatory Responses of Human Airway Cells to Ricin Involve Stress-Activated Protein Kinases and NF-kB              |

These data are downloaded from NCBI GEO web site. <http://www.ncbi.nlm.nih.gov/geo/>

## Supplementary Data (C): Data source detail for Mouse.

| GSE ID  | Number of slides | Experimental type                      | Sample description  |
|---------|------------------|--|---|
| GSE1074 | 12               | Ordered                                | Strain-dependent effects of alcohol on early mouse embryos  |
| GSE1435 | 27               | other                                  | Microarray Based Comparison of two Amplification Methods For Nanogram Amounts of Total RNA                            |
| GSE1479 | 36               | time-course                            | C57BL/6 Benchmark Set for Early Cardiac Development   |
| GSE1835 | 4                | dose response                          | Myocilin effect on retina. Russell-7U24NS043571-02  |
| GSE1871 | 12               | other                                  | SCCOR_MouseLung_Simva_LPS   |
| GSE1986 | 17               | other                                  | Global gene expression across a range of tissues  |
| GSE1999 | 15               | dose response                          | Neuroprotective effects of erythropoietin. Wilson-5R01NS028208-15   |
| GSE2019 | 11               | ordered                                | Microarray Based Comparison of three Amplification Methods For Nanogram Amounts of Total RNA                          |
| GSE2060 | 5                | parallel sample                        | Characterize CREB target genes in different tissue types  |
| GSE2161 | 8                | other                                  | Identification of genes that are dysregulated in the telencephalon of Dlx1/2 mutants. Rubenstein-2R01MH049428-11      |
| GSE2372 | 6                | other                                  | Aortae of 32 weeks old apoE mice  |
| GSE2389 | 4                | repeat sample                          | Regulatory T cell lineage specification by Foxp3  |
| GSE2463 | 2                | ordered                                | Targets of FGFR2-IIIb signalling in the hair follicle   |
| GSE2843 | 5                | treatment response                     | thymic mouse cells  |
| GSE2869 | 8                | other                                  | Rosenfeld-5R01NS034934-14   |
| GSE2873 | 4                | other                                  | Burden-2R01NS036193-06A1  |
| GSE3100 | 23               | disease state analysis                 | Cystic Fibrosis Mouse Lung Profiles   |
| GSE3203 | 16               | Infection-induced responses            | Influenza virus infection-induced gene expression changes of regional B cells are mediated in part through type I IFN |
| GSE3414 | 36               | Time Course                            | Immune Response to Nippostrongylus brasiliensis in the mouse lung   |
| GSE3425 | 2                | treatment response                     | Silencing of microRNAs in vivo with "antagomirs"  |
| GSE3440 | 14               | time series                            | Effect of aldosterone on gene expression in the heart   |
| GSE3463 | 12               | Time course and comparison of sex      | Expression profiling of mouse gonadal somatic cells   |
| GSE3554 | 6                | retina, glaucoma, DBA/2J, elevated     | Microarray Analysis of Retinal Gene Expression in the DBA/2J Model of Glaucoma  |
| GSE3653 | 16               | Ngn3 induction, embryonic stem ce      | Inducible Ngn3 Embryonic Stem Cells   |
| GSE3822 | 16               | embryonic metanephric kidney, kidn     | Complete embryonic E11.5 metanephric kidney analysis of wild type and Hoxa11, Hoxd11 null targeted mice               |
| GSE4034 | 24               | dose response                          | palme-afly-mouse-198967   |
| GSE4035 | 24               | NIH Neuroscience Microarray Cons       | palme-afly-mouse-84746  |
| GSE4040 | 6                | dose response                          | sikes-afly-mouse-84486  |
| GSE4041 | 6                | other                                  | zhang-afly-mouse-200705   |
| GSE4043 | 6                | time course, cell type comparison      | Gene profiling analysis of Src chemical rescue  |
| GSE4051 | 39               | Transcription factor, development,     | Targeting of GFP to new-born rods by Nr1 promoter and temporal expression profiling of flow-sorted photoreceptors     |
| GSE4098 | 16               | repeat                                 | Mouse limbal epithelial basal cells vs. corneal epithelial basal cell gene expression profile                         |
| GSE4142 | 14               | Cell Type Comparison                   | Molecular Analysis of antigen-specific B cell responses   |
| GSE4179 | 3                | cell type comparison                   | A function for interleukin 2 in Foxp3-expressing regulatory T cells   |
| GSE4189 | 14               | Array Based                            | The Oct4 and Nanog transcription network that regulates pluripotency in mouse embryonic stem cells                    |
| GSE4193 | 8                | Specific germ cells from murine tes    | Gene expression in murine germ cells  |
| GSE4230 | 8                | tissue specificity, time course, devt  | Gene expression profiles in developing nephrons using Lim1 metanephric mesenchyme-specific conditional mutant mice    |
| GSE4260 | 6                | time course                            | Cumulus-oocyte complex temporal expression  |
| GSE4307 | 20               | Single cell analysis                   | Expression data from single cells from ICMs of mouse blastocysts at E3.5  |
| GSE4308 | 16               | Method validation                      | Expression data for validation of single cell cDNA amplification method (V1V3 method)                                 |
| GSE4309 | 36               | SuperSeries                            | An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis      |
| GSE4332 | 8                | time_series_design                     | Cell intrinsic alterations underlie hematopoietic stem cell aging   |
| GSE4411 | 11               | other                                  | burde-afly-arabi-64764  |
| GSE4413 | 4                | other                                  | katz-afly-mouse-212743  |
| GSE4481 | 18               | Growth plate zone comparison and       | C-type natriuretic peptide regulates endochondral ossification through p38 MAP kinase-dependent pathways_2            |
| GSE4658 | 6                | static vs simulated microgravity       | static vs simulated microgravity  |
| GSE4695 | 6                | endothelin-1; Et1; dermal fibroblast   | Changes in gene expression in dermal fibroblasts following exposure to Et1 peptide                                    |
| GSE4752 | 6                | other                                  | tourt-afly-mouse-243760   |
| GSE4756 | 8                | dose response                          | ruiz-afly-mouse-243540  |
| GSE4758 | 24               | time-course                            | conra-afly-mouse-89474  |
| GSE4774 | 15               | other                                  | ruben-afly-mouse-187820   |
| GSE4786 | 9                | Effect of aging, effect of caloric res | Caloric restriction suppresses apoptotic cell death in the mammalian cochlea and leads to prevention of presbycusis   |
| GSE4866 | 10               | disease state analysis                 | The role of mtDNA mutations in age-related hearing loss in mice carrying a mutator DNA polymerase gamma               |
| GSE4936 | 8                | neural progenitors, embryoid bodies    | Transcriptional Profiling of Hh-Treated Embryoid Bodies   |
| GSE5018 | 3                | Bhlhb8, Mist1, mucous neck cell, lar   | A profile of murine gastric epithelial cells: Parietal, Zymogenic, Pit  |
| GSE5035 | 12               | time-course                            | zhang-afly-arabi-50806  |
| GSE5036 | 4                | dose response                          | zhang-afly-mouse-175010   |
| GSE5037 | 18               | dose response                          | zhang-afly-mouse-217286   |
| GSE5038 | 9                | other                                  | zhang-afly-mouse-308606   |
| GSE5124 | 6                | gene swap, knock-in, hindbrain devt    | Hoxb1 gene expressing Hox-A1 protein: Altered transcriptional profile in rhombomere 4 at E10.5                        |
| GSE5127 | 18               | toxicology, chemical carcinogenesis    | Gene Expression Biomarkers for Predicting Lung Tumors in Two-Year Rodent Bioassays                                    |
| GSE5128 | 18               | toxicology, chemical carcinogenesis    | Gene Expression Biomarkers for Predicting Liver Tumors in Two-Year Rodent Bioassays                                   |
| GSE5156 | 4                | functional genomics; transcriptional   | Impact of intestinal colonization on Paneth cell gene expression  |
| GSE5198 | 18               | response to microbial colonization     | Transcriptional profiling of mouse ileum in response to colonization with a zebrafish or mouse gut microbiota         |
| GSE5202 | 12               | time course                            | Expression data from B. melitensis infected mouse macrophages   |
| GSE5296 | 96               | Time-Series                            | Spinal Cord Injury Murine Model   |
| GSE5298 | 8                | genetic modification                   | Development of heart valves requires Gata4 expression in endothelial-derived cells                                    |
| GSE5309 | 7                | Normal C57BL/6 Mammary Gland E         | Transcriptional Profiling of Mammary Gland Side Population Cells  |
| GSE5324 | 48               | mRNA degradation, time course, kn      | TTP mRNA targets identified by global analysis of stabilized transcripts in TTP-deficient fibroblasts                 |
| GSE5338 | 8                | genetic modification                   | In vivo function of NR2E3 in establishing photoreceptor identity during mammalian retinal development                 |
| GSE5500 | 21               | genetic modification, pressure over    | Gata4 is required for postnatal cardiac function and protection from pressure overload-induced heart failure          |
| GSE5582 | 6                | neocortical astrocyte, cerebellar as   | Neocortical and cerebellar astrocytes   |
| GSE5587 | 6                | other                                  | tourt-afly-arabi-307860   |
| GSE5657 | 20               | 2x2 factorial design with complete     | Effect of X-linked hypophosphatemia (the Hyp mutation) on gene expression in the mid-shaft of the mouse femur         |
| GSE5668 | 6                | Developmental stage (mature vs. in     | Identification and characterization of the changed and stable transcripts during mouse oocyte maturation              |
| GSE5715 | 10               | disease state analysis                 | Intestinal Phenotype of Variable Weight Cystic Fibrosis Knockout Mice   |
| GSE5817 | 21               | other                                  | marsh-afly-mouse-232749   |
| GSE5959 | 6                | Genetic variation                      | Expression differences in the liver of a congenic mouse with low serum IGF-1  |
| GSE5976 | 12               | time course                            | Mesp1-induced gene expression changes   |
| GSE5985 | 4                | cell stimulation                       | Gene expression profile of BAFF-stimulated B cells  |
| GSE5987 | 11               | Tumor vs normal comparisons            | Mouse model of Endometrioid Ovarian Adenocarcinomas by conditional inactivation of PTEN and APC genes                 |



|         |     |                                      |  |
|---------|-----|--------------------------------------|--|
| GSE6055 | 8   | disease state analysis, time course  | Gene Expression Profiling Reveals Unique Pathways Associated with Differential Severity of Lyme Arthritis                        |
| GSE6065 | 100 | Aeromonas; Virulence; Gene expres    | Murine host cell response to Aeromonas infection   |
| GSE6078 | 5   | Disease state analysis, genetic mod  | PTEN-deficient intestinal stem cells initiate intestinal polyposis   |
| GSE6079 | 2   | expression comparison                | expression comparison of wildtype lungs and hypomorphic for nmyc   |
| GSE6116 | 70  | toxicology, chemical carcinogenesis  | Transcriptional Biomarkers to Predict Female Mouse Lung Tumors in Rodent Cancer Bioassays - A 13 Chemical Training Set           |
| GSE6210 | 12  | Liver and quadriceps muscle gene e   | Hypomorphic Mutation in PGC1beta causes mitochondrial dysfunction and liver insulin resistance                                   |
| GSE6223 | 13  | Gene expression comparison from r    | Gene expression profiles of embryonic day 14 bladder, urogenital sinus, and urethra.   |
| GSE6257 | 2   | TNFalpha, Lymphatic endothelium, i   | An inflammation-induced mechanism for leukocyte transmigration of lymphatic vessel endothelium.                                  |
| GSE6259 | 21  | cell type comparison of wildtype an  | Differential antigen processing by dendritic cell subsets in vivo  |
| GSE6290 | 37  | Comparison of kidney components.     | Gene expression profiles of components of the developing kidney  |
| GSE6291 | 6   | mRNA expression profiling, oligonuc  | Transcriptome Analysis Multipotent Adult Progenitor Cells (Affy)   |
| GSE6376 | 4   | stress response                      | Zymosan stimulation of MyD88-deficient bone marrow-derived macrophages   |
| GSE6383 | 6   | comparative genomic hybridization:   | Mouse small intestine epithelium vs. mesenchyme  |
| GSE6397 | 14  | disease state modification           | Comparison between gene expression in heart from Lmna H222P heterozygous and control mice  |
| GSE6398 | 14  | disease state modification           | Comparison between gene expression in heart from Lmna H222P homozygous and control mice  |
| GSE6399 | 14  | disease state modification           | Comparison between gene expression in heart from Emd KO and control mice   |
| GSE6424 | 4   | hematopoietic stem cells, immortal   | B-catenin Plays a Dual Role in the Generation and Maintenance of Immortalized Hematopoietic Stem Cells                           |
| GSE6461 | 9   | tumor versus normal muscle tissue    | SYT-SSX murine synovial sarcoma model  |
| GSE6466 | 7   | Gene expression comparison betwe     | Gene expression profiles of postnatal day 1 and adult bladders   |
| GSE6482 | 9   | Cell type comparison                 | mECK36: a cell and animal model of virally induced Kaposi's sarcoma  |
| GSE6487 | 30  | Time-Series                          | Myogenesis MyoD  |
| GSE6514 | 90  | gene expression, temporal changes.   | Gene expression in the mouse brain during spontaneous sleep and prolonged wakefulness  |
| GSE6526 | 16  | time course                          | Expression time course data HEY2 KO and WT MAMC treated with PDGF  |
| GSE6589 | 11  | Comparison of kidney components.     | Gene expression profiles of components of the developing kidney (GUDMAP Series_id: 12)   |
| GSE6595 | 40  | embryo kind comparison               | Gene expression study in 1 and 2 cell cloned mice embryos  |
| GSE6623 | 12  | Hematopoietic stem cells, myeloid ;  | FoxO are critical mediators of hematopoietic stem cell resistance to physiologic oxidative stress                                |
| GSE6675 | 8   | fibroblast growth factor-2 treatme   | Astroglial gene expression program elicited by fibroblast growth factor-2 mande-ffy-mouse-307080                                 |
| GSE6676 | 8   | TGF-beta overexpression              | Comparison of corneas from wildtype mice to those from mice under the influence of high doses of TGF-beta russe-ffy-mouse-372894 |
| GSE6678 | 18  | gene knockout                        | Palmitoyl protein thioesterase-1 knockout mice hofma-ffy-mouse-370000  |
| GSE6681 | 4   | genetic modification                 | Foxp3 ablation in peripheral mature regulatory T cells   |
| GSE6765 | 9   | Aeromonas caviae, infection, diseas  | Aeromonas caviae infection, 24 hours   |
| GSE6770 | 4   | Genetic Modification                 | Gene Expression Data in HDAC2 KO Myocardium  |
| GSE6789 | 12  | genetic modification, strain compari | response to IL-1b of WT and IRAK4 kinase dead mouse embryonic fibroblasts  |
| GSE6815 | 4   | Melanotransferrin, hyperexpression   | Hyperexpression of Mouse Melanotransferrin on LMTK Cell Line   |
| GSE6833 | 4   | SuperSeries                          | Hyperexpression and Downregulation of Melanotransferrin on Various Cell Lines  |
| GSE6837 | 8   | comparative study                    | Expression data from wild type (wt) and lkbke knockout (lke) embryonic fibroblasts (EF)  |
| GSE6858 | 16  | case control                         | Expression data from experimental murine asthma  |
| GSE6867 | 6   | genetic modification                 | Expression data in the absence of Notch1 in hair follicles   |
| GSE6875 | 8   | Cell type comparison                 | Development of Regulatory T cell Precursors in the Absence of a Functional Foxp3 Protein   |
| GSE6881 | 10  | Time Course                          | Embryonic Testis Developmental Time Course   |
| GSE6882 | 10  | Time course                          | Embryonic Ovary Developmental Time Course  |
| GSE6916 | 20  | SuperSeries                          | Embryonic Testis/Ovary Developmental Time Courses  |
| GSE6934 | 4   | mutant comparison                    | Wnt4-dependent gene expression in the developing kidney (GUDMAP Series_id: 13)   |
| GSE6945 | 4   | mouse embryonic fibroblasts, gene    | The Retinoblastoma Binding Protein RBP2 is an H3K4 demethylase   |
| GSE6958 | 2   | cell treatment comparison            | FOG-1-independent transcription by GATA-1(V205G) in G1E cells  |
| GSE6959 | 12  | repeat                               | Oscillating gene expression in the prefrontal cortex of C57Bl/6J mice at Zeitgeber Time (ZT) 3, 9, 15, and 21.                   |
| GSE6963 | 4   | Gene expression analysis of the po   | Gene expression profiles of mouse postnatal day 1 ureter (GUDMAP Series_id: 14)  |
| GSE6980 | 16  | XBP-1, MGUS, multiple myeloma, ti    | The Differentiation and Stress Response Factor, XBP-1, Drives Multiple Myeloma Pathogenesis                                      |
| GSE6998 | 32  | time course                          | Expression profiling of developmental and regenerating liver in mice   |
| GSE7012 | 13  | time series                          | Identif. of oscillatory genes in somitogenesis from functional genomic analysis of C2C12 myoblast line                           |
| GSE7020 | 8   | strain comparison                    | The molecular consequences of Nix ablation on apoptosis and erythropoiesis   |
| GSE7050 | 18  | Lckcre, CD4Cre-Ctnnbx3, lymphoi      | Stabilization of b-catenin induces lymphomas   |
| GSE7054 | 13  | SuperSeries                          | Identification of oscillatory genes in somitogenesis from functional genomic analysis  |
| GSE7069 | 8   | Global gene expression data analys   | Zfx controls the self-renewal of embryonic and hematopoietic stem cells  |
| GSE7196 | 6   | microarray and genetic modificatio   | Differential gene expression between WT and ERRA-null hearts   |
| GSE7257 | 9   | Lim1 mutant, kidney development, e   | Laser capture-microarray analysis of Lim1 mutant kidney development.   |
| GSE7280 | 10  | cell type comparison, response to e  | Gene expression in peripheral cells: effects of Foxp3 and PDE3B  |
| GSE7302 | 6   | cell type comparison                 | Expression data from bone marrow hematopoietic stem cell CD34 Flt3 subfractions  |
| GSE7324 | 10  | genetic modification                 | Structural basis for recognition of SMRT/N-CoR by the MYND domain and its contribution to AML1/ETO's activity                    |
| GSE7333 | 6   | miRNA                                | miR-1-2 knockout versus wild-type hearts   |
| GSE7342 | 12  | time course                          | Expression data from p38 knock out versus wild type fetal liver  |
| GSE7348 | 6   | Treatment Comparison                 | Gene Expression in Naive and Tolerant Macrophages stimulated with LPS  |
| GSE7381 | 6   | knockout and wildtype mice compar    | Expression profiling of back skin of Get1 and control mice at e18.5  |
| GSE7404 | 127 | time course                          | Comparison of Longitudinal Leukocyte Gene Expression after Burn Injury or Trauma Hemorrhage in Mice                              |
| GSE7407 | 4   | Gene expression comparison           | Regulation of Gene Expression by Sirt1 in the heart  |
| GSE7424 | 8   | transplantation tolerance            | Intra-graft gene expression profile associated with the induction of tolerance   |
| GSE7430 | 12  | Time course                          | Investigate the role of Wnt/beta-catenin signaling in development of the exocrine pancreas                                       |
| GSE7764 | 10  | mRNA, mouse, NK cell, IL-15          | mRNA expression profiles of resting and IL-15 activated murine NK cells  |
| GSE7770 | 4   | cell type comparison, development    | Gene expression in thymic CD4 T cells: effects of Foxp3  |
| GSE7773 | 14  | SuperSeries                          | Foxp3-dependent programme of regulatory T-cell differentiation   |
| GSE7784 | 12  | developmental stages                 | Expression profiling of double negative thymocytes   |
| GSE7810 | 8   | comparative expression profiling     | Comparative analysis of gene expression WT and Nrf2-/- mice Type II cells  |
| GSE7823 | 9   | Differential response to hypoxia by  | Murine Pulmonary Response to Chronic Hypoxia is Strain Specific  |

These data are downloaded from NCBI GEO web site. <http://www.ncbi.nlm.nih.gov/geo/>

## Supplementary Data (D): Data source detail for Rat.

| GSE ID  | Number of slides | Experimental type                     | Sample description  |
|---------|------------------|---------------------------------------|---|
| GSE2870 | 18               | dose response                         | Ogle-5P01NS037520-05  |
| GSE2872 | 12               | dose response                         | d'mel-afly-rat-168311   |
| GSE2875 | 6                | dose response                         | Sabban-5R01NS028869-11  |
| GSE3298 | 16               | time course, fracture, growthplate,   | mRNA expression in rat proximal femoral growthplate after mid-shaft fracture  |
| GSE3866 | 10               | Gene transfer                         | p38 MAPK overexpression in heart in vivo.   |
| GSE4031 | 105              | dose response                         | colan-afly-rat-89421  |
| GSE4037 | 18               | other                                 | prosc-afly-rat-185420   |
| GSE4038 | 4                | dose response                         | sabba-afly-rat-168529   |
| GSE4753 | 40               | time-course                           | wilso-afly-rat-132990   |
| GSE4776 | 83               | other                                 | Thrivikraman-5P50MH058922-050002  |
| GSE4865 | 6                | rat pituitary organ culture treated v | Comparison of gene expression profiles between GnRH-treated and control rat pituitary organ cultures                |
| GSE5034 | 8                | time-course                           | Thompson-5R01NS040338-02  |
| GSE5282 | 11               | rat optic nerve head astrocytes       | Astrocytes in response to activation of EGFR  |
| GSE5350 | 72               | Cross-platform comparison             | MicroArray Quality Control (MAQC) Project   |
| GSE5509 | 39               | other                                 | Expression data from Rat liver 48 hours after treated with different toxic compounds.                               |
| GSE5606 | 14               | single time point, comparison contr   | Expression data from Normal (control) and Diabetic animals LV heart tissue (At 16wks)                               |
| GSE5680 | 120              | eQTL, F2 cross                        | Expression data for an eQTL experiment in rat eye   |
| GSE6497 | 15               | gene expression changes due to aci    | Expression profile of syngeneic (sTX) and allogeneic kidney (aTX) transplantation compared to control (ctr) kidneys |
| GSE7323 | 9                | dose response                         | Hippocampus response to ketone bodies (maalo-afly-rat-387180)   |
| GSE7491 | 9                | Time course                           | Expression data from rat lung alveolar development  |
| GSE7493 | 9                | time-course                           | Mutant SOD1 rats (lobsi-afly-rat-194438)  |
| GSE7697 | 8                | Dose response                         | Granulosa cell gene expression associated with oocyte developmental competence                                      |

These data are downloaded from NCBI GEO web site. <http://www.ncbi.nlm.nih.gov/geo/>