Supplementary Appendix

Potential Therapeutic Targets for Lung Repair During Human Ex Vivo Lung Perfusion

Aaron Wong[1], BSc., Ricardo Zamel[2], PhD., Jonathan Yeung[2,3], MD, PhD., Gary D. Bader[4], PhD., Claudia C Dos Santos[1,5], MD, MSc., Xiaohui Bai[2], MSc., Yubo Wang[2], Shaf Keshavjee[1,2,3]*, MD, MSc., Mingyao Liu[1,2,3]*, MD, MSc.

*These authors share senior authorship.

Affiliations:

1. Institute of Medical Science, Faculty of Medicine, University of Toronto
2. Latner Thoracic Surgical Research Laboratories, Toronto General Hospital Research Institute, University Health Network
3. Toronto Lung Transplant Program, Department of Surgery, Faculty of Medicine, University of Toronto
4. Molecular Genetics, Faculty of Medicine, University of Toronto
5. Keenan Research Centre for Biomedical Science, St. Michael's Hospital

Supplementary Appendix: Table of Contents

Content	Page
Supplementary Methods	1-2
Supplementary Figure S1. Comparison of transplantation and EVLP as Models of Reperfusion.	3
Supplementary Figure S2. Transplant Validation - Gene Set Clusters Enriched from the Kang et al. Dataset.	4
Supplementary Figure S3. EVLP Validation - Gene Set Clusters Enriched from the Yeung et al. Dataset.	5
Supplementary Figure S4. Cluster classification across a range of predominance thresholds between 0.5-1.0.	6
Supplementary Table S1. Enriched Gene Set Clusters from Transplant and EVLP Comparison.	7-8
Supplementary Table S2. Number of FDA approved therapeutics predicted to target pathways	9-10
Supplementary Data 1. Comparison of young and old donor lungs within transplant and EVLP datasets.	11
References	12

Supplementary Methods:

RNA Extraction and Microarrays. Peripheral lung tissue biopsies were collected and snap frozen in liquid nitrogen in the operating room. Total RNA was purified using RNeasy Mini Kit (Qiagen; Hilden, Germany). RNA quality was assessed using Nanodrop spectrophotometer (VWR; Radnor, PA) and Bioanalyzer (Agilent; Santa Clara, CA). Samples with concentration above 100ng/ul and RNA Intergaty Number (RIN) above 7.0 were used for microarray analysis. Purified RNA was stored at -80°C. Microarrays were processed by Princess Margret Genomics Center (Toronto, Canada). Transplant samples had gene expression measured on Human Genome U133 Plus 2.0 Array (Affymetrix; Santa Clara, CA) and EVLP samples on Clariom D Assay (Affymetrix). Analysis was conducted in R version 3.5.1 (1).

Microarray Preprocessing. Microarrays were normalized using Robust Multi-array Average (RMA) (2) in the oligo package (3). Genes were annotated using Brainarray version 22 custom annotation files (4).

Limma Analysis. Differential gene expression (DGE) was calculated using Bayes moderated paired t-test in the Limma package (5). The transplant group had differential gene expression calculated between paired post and pre transplant samples, while the EVLP group had differential expression calculated between paired post and pre EVLP samples. Differentially expressed genes were defined as having an FDR<0.05 using the Benjamin-Hochberg correction procedure (6).

Generation of Ranked list for Gene Set Enrichment Analysis. Only the genes detectable on both microarray platforms were used in the generation of ranked lists. For each dataset, genes were ranked based on their score according to the formula:

$$\text{Gene score} = -\ln(\text{p value from DGE})(\text{Sign of gene fold change})$$

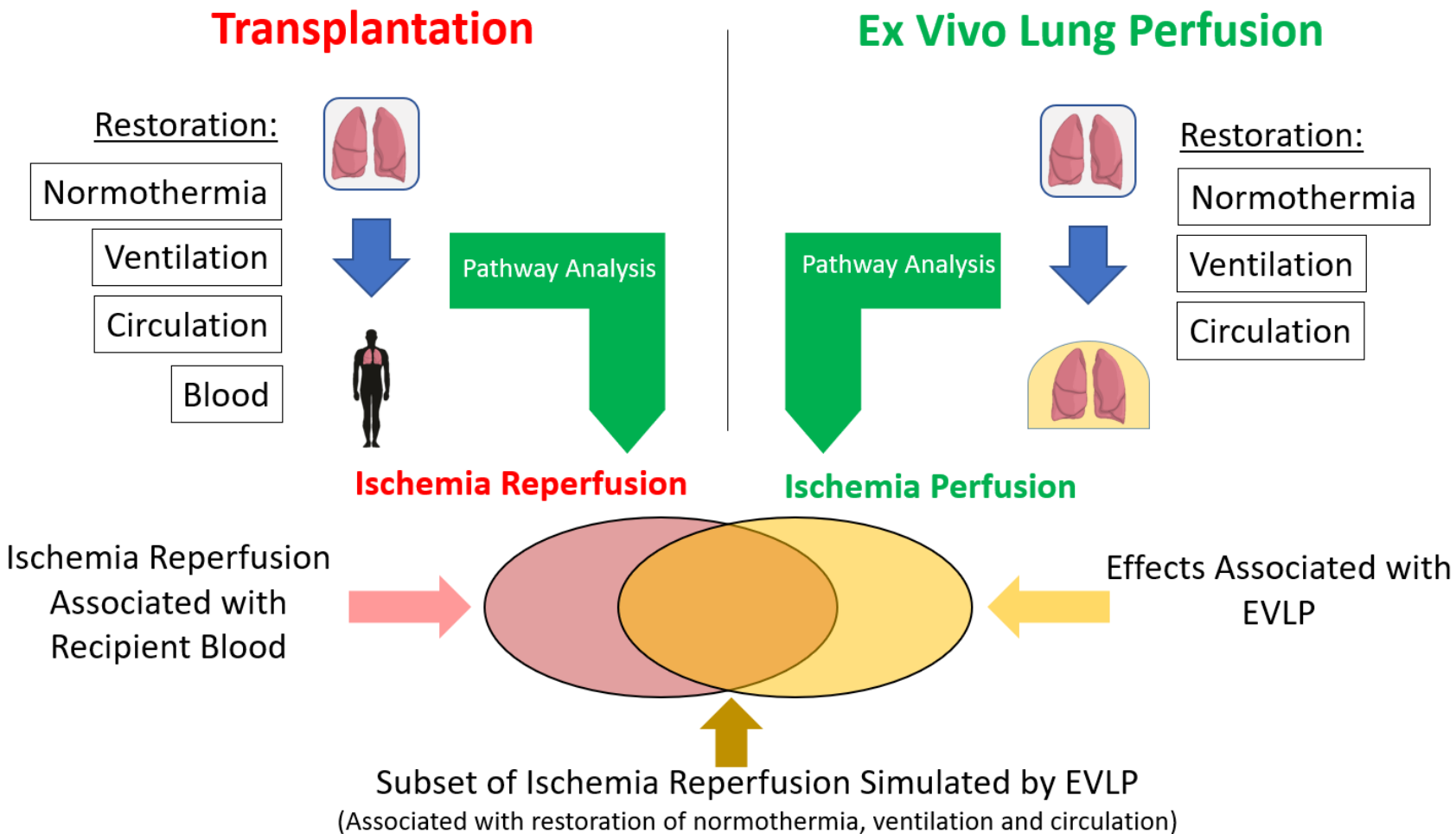
Ranked lists were then passed to GSEA.

Gene Set Enrichment Analysis 3.0 (GSEA) (7)**.** A PreRanked analysis was run with the following parameters: gene set min=15, gene set max=500, permutations=1000, scoring scheme = weighted, normalization mode = mean div. The gene set database used is from the Bader lab: Human_GOBP_AllPathways_no_GO_iea_June_01_2017_entrezgene.gmt from the Bader Lab. More information on the database can be found at: http://baderlab.org/GeneSets.

Cytoscape Visualization. Gene set reports were filtered for gene sets which met the FDR < 0.05. Cytoscape version 3.5.0 was used (8). Pathway network was generated using EnrichmentMap version 2.2.1 (9) using the follow parameters: similarity overlap = overlap coefficient, cutoff=0.5. Clusters were annotated with AutoAnnotate 1.1.0 (10) using the parameters: clustering algorithm = MCL, edge weight column = similarity coefficient. Each cluster with more than 4 nodes was reviewed and had a name assigned.

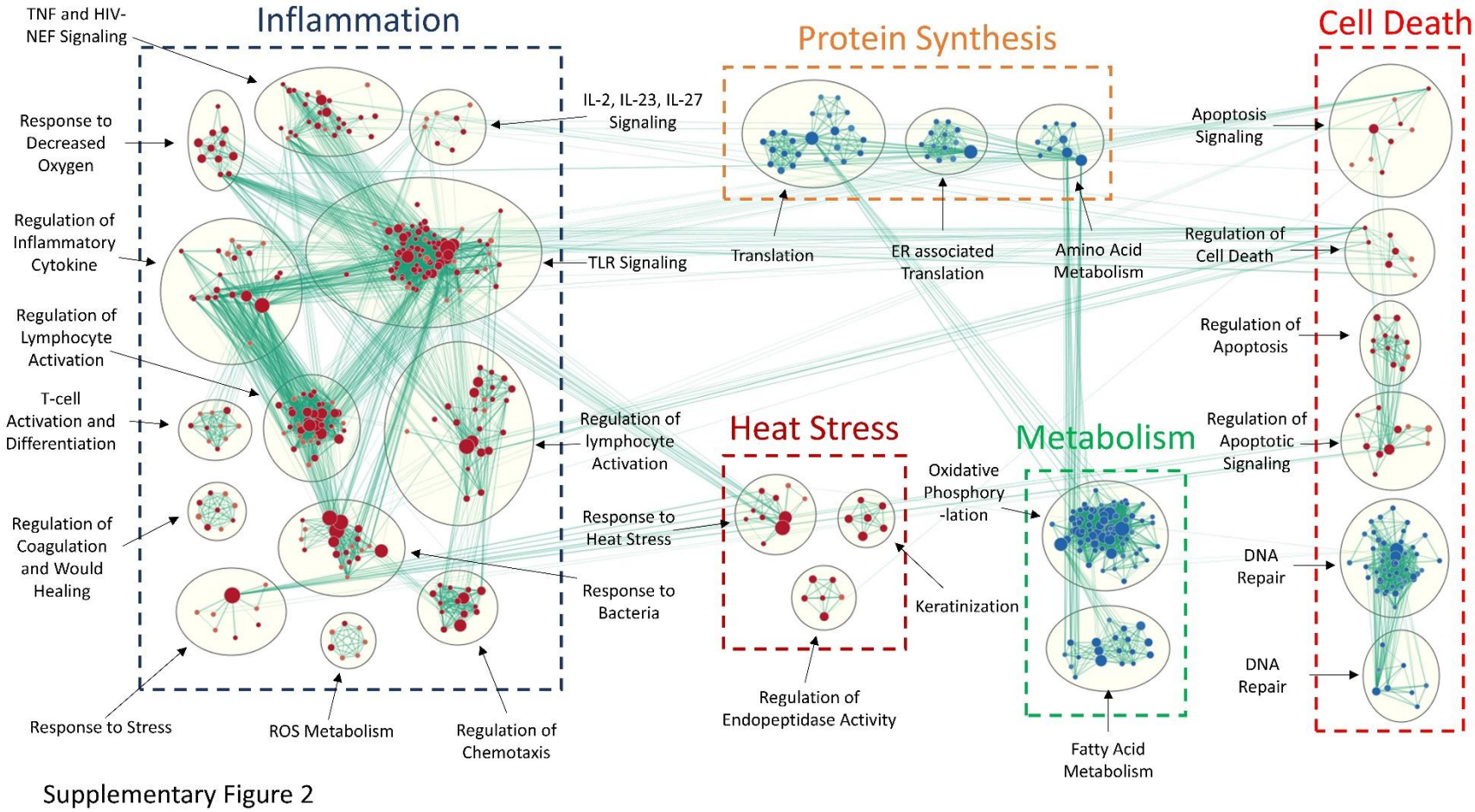
PCA Visualization. PCA visualization were made using the R package pca3d (11). All genes from each microarray platform were used to create plots.

Validation Analysis. Validation datasets were analyzed using the same pipeline as the study dataset. Due to the small sample size of both datasets, we increased the cutoffs for gene sets to FDR<0.10.



Supplementary Figure 1. Comparison of Transplantation and EVLP as Models of Reperfusion.

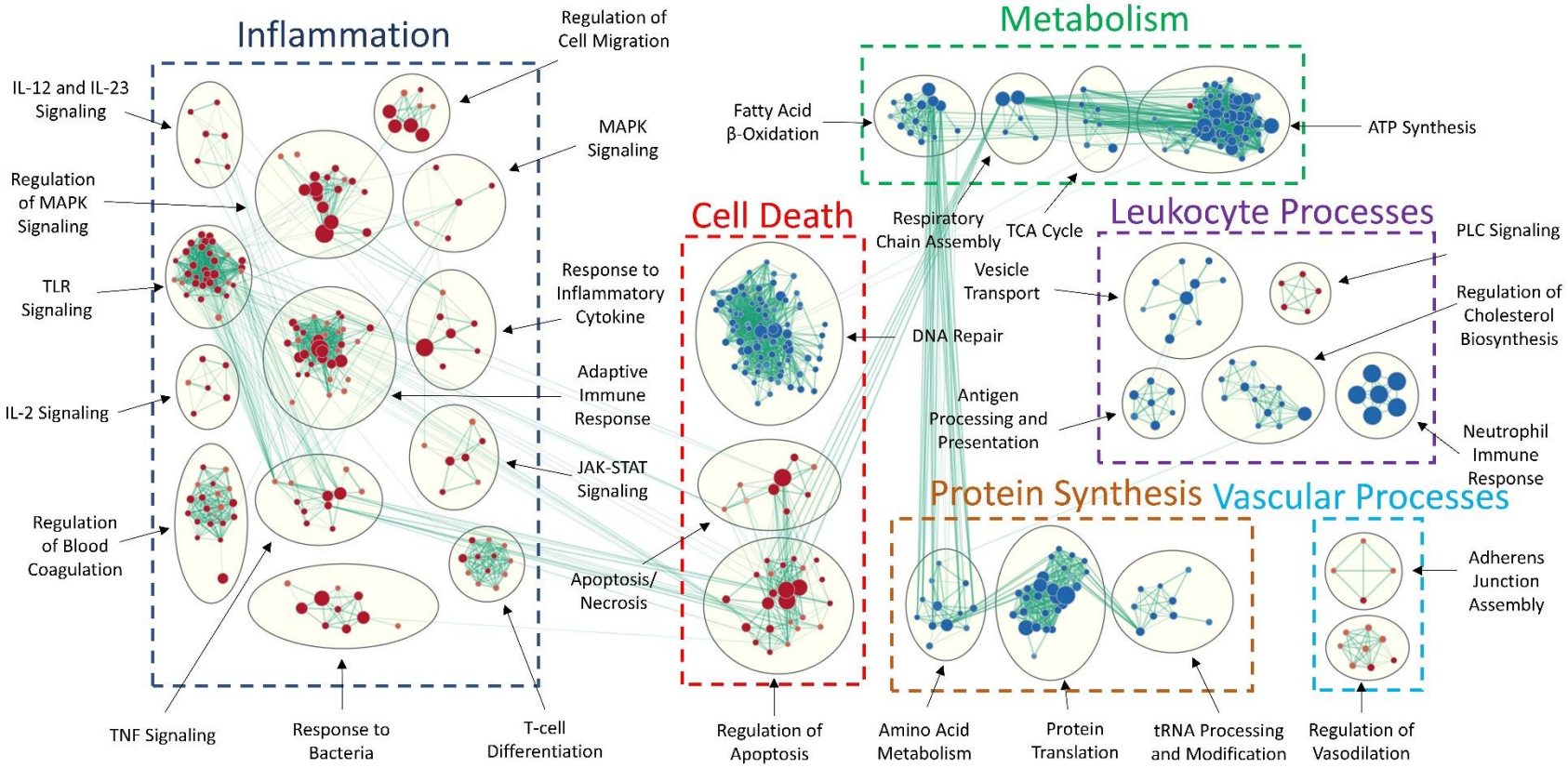
The left side of the panel represents the biological response induced by recipient reperfusion, while the right side represents the response induced by EVLP. The Venn diagram represents a comparison of the enriched biological pathways in transplantation and EVLP, which can be divided into three categories.



Supplementary Figure 2

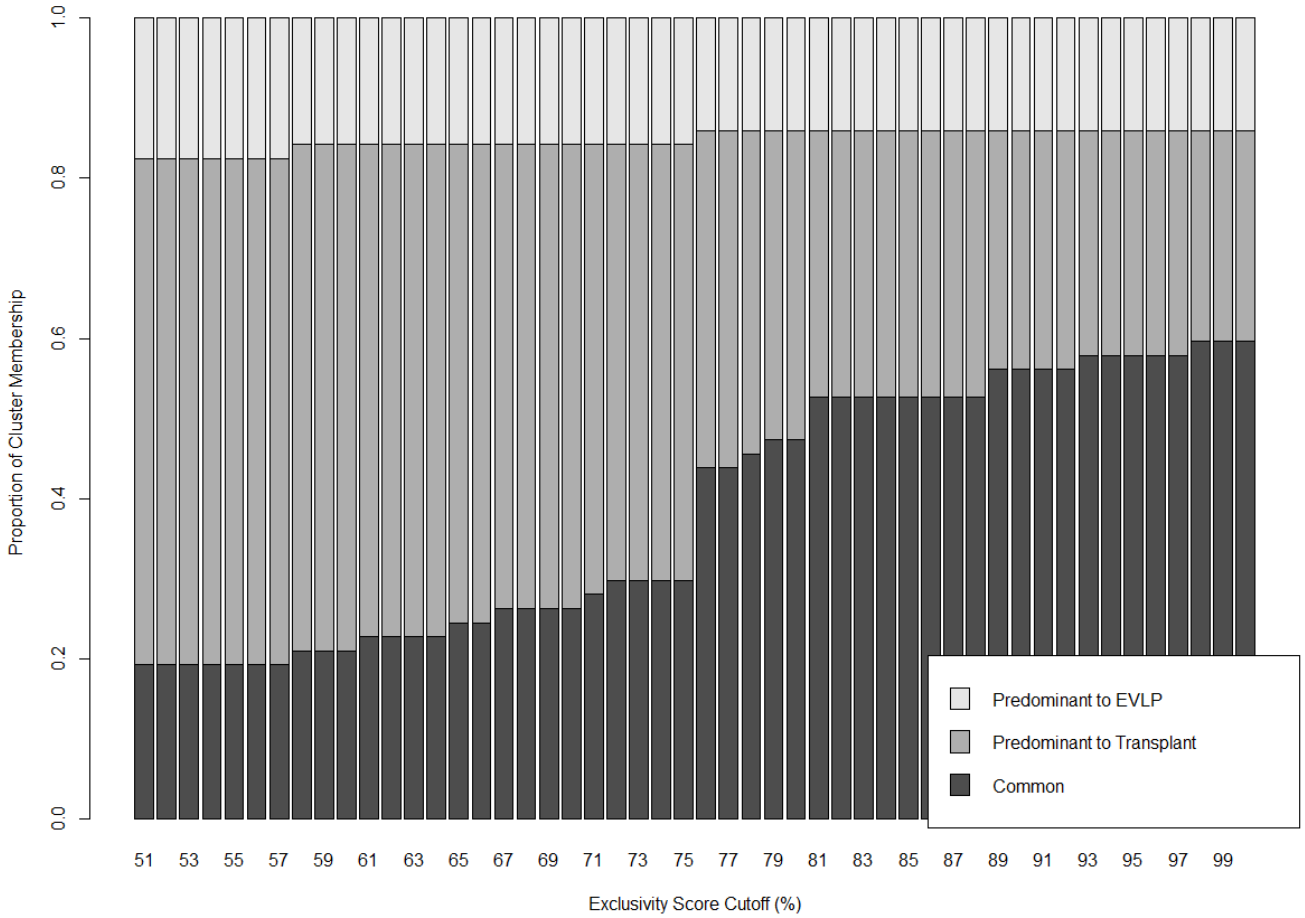
Supplementary Figure 2. Transplant Validation - Gene Set Clusters Enriched from the Kang et al. Dataset (12)**.**

Gene set clusters which met cutoffs (FDR<0.10) fell into five major categories with inflammation, cell death and heat stress were generally up-regulated (red nodes) and metabolism and protein synthesis down-regulated (blue nodes). The gene set clusters observed here recapitulate the same themes classified as being predominant to transplant and common from the study dataset.



Supplementary Figure 3. EVLP Validation - Gene Set Clusters Enriched from the Yeung et al. Dataset(13)**.**

Gene set clusters which met cutoffs (FDR<0.10) fell into six major categories with inflammation, cell death and vascular processes generally up-regulated (red nodes) and metabolism, protein synthesis and leukocyte processes down-regulated (blue nodes). The gene set clusters observed here recapitulate the same themes classified as being predominant to EVLP and common from the study dataset.



Supplementary Figure 4. Proportion of cluster membership across a range of thresholds between 0.5-1.0.

Supplementary Table S1. Enriched Gene Set Clusters from Transplant and EVLP Comparison |0.8| threshold.

Cluster Name	Number of Gene Sets	Predominance Score	Classification
TLR/MYD88 Signaling	99	0.71	common
Regulation of Adaptive Immunity	58	0.76	common
Translation	49	0.78	common
Response to Bacteria	46	0.72	common
Regulation of Apoptic Signaling	28	0.64	common
Cillium Organization	14	0.29	common
Regulation of Leukocyte Chemotaxis	14	0.79	common
Fatty Acid Beta Oxidation	14	0.43	common
Regulation of Blood Coagulation	13	0.23	common
Response to TNF and IL-1	13	0.23	common
Amino Acid Metabolism	12	0.75	common
DNA Repair	11	0.18	common
Epigenetic Regulation of Expression	11	0.09	common
IL-12 and IL-23 Signaling	9	0.67	common
Regulation of Alternative Splicing	8	0.75	common
S1P Signaling	8	0.75	common
Keratinization	7	-0.57	common
unknown	5	0.20	common
Tubulin Folding	5	-0.40	common
Response to Hormone	5	0.40	common
Cofactor Metabolic Process	5	0.60	common
Amino Acid Metabolism	4	0.25	common
Intracellular Transport	4	-0.75	common
AP1 and Fra Pathway	4	0.00	common
unknown	4	0.75	common
Regulation of Endocrine Process	4	0.75	common
NFAT and TCR Pathway	4	0.75	common
Oxidative Phosphorylation	74	0.97	transplant predominant
Regulation of MAPK Signaling	26	0.81	transplant predominant
HIV-NEF Signaling and TNF Signaling	26	0.81	transplant predominant
Leukocyte Chemotaxis	17	0.88	transplant predominant
Cell Death Signaling	13	0.92	transplant predominant
IL-2 and GMCSF Signaling	12	1.00	transplant predominant
TCR, BCR, IL-2 Signaling	10	0.80	transplant predominant
Regulation of Protein Import to Nucleus	9	0.89	transplant predominant
Regulation of FGFR Signaling	9	1.00	transplant predominant
DNA Repair-1	8	1.00	transplant predominant
DNA Repair-2	8	1.00	transplant predominant
Regulation of ROS Metabolism/Biosynthesis	7	1.00	transplant predominant
Regulation of JAK-STAT Signaling	7	1.00	transplant predominant
MAPK Signaling	6	1.00	transplant predominant

Symbiont Host Modulation	6	1.00	transplant predominant
ERBB and TRK Receptor Signaling	5	1.00	transplant predominant
Response to Inflammatory Cytokine	5	1.00	transplant predominant
TCR and BCR Signal Transduction	5	1.00	transplant predominant
Leukocyte Mediated Cytotoxicity	5	1.00	transplant predominant
Leukocyte Homeostasis	4	1.00	transplant predominant
Thromboxane and IL-8 Signaling	4	1.00	transplant predominant
Heart Morphogenesis	4	1.00	transplant predominant
Phosphatidylinositol Biosynthesis	17	-1.00	EVLP predominant
PLC Signaling	9	-1.00	EVLP predominant
Vesicle Transport	6	-1.00	EVLP predominant
Regulation of Vasodilation	6	-1.00	EVLP predominant
Adherens Junctions	5	-1.00	EVLP predominant
Protein Localization to Vacuole	5	-1.00	EVLP predominant
Cholesterol Biosynthesis	4	-1.00	EVLP predominant
Regulation of Epithelial and Endothelial Apoptosis	4	-1.00	EVLP predominant

*Red cluster names indicate upregulation while blue names indicate downregulation of the pathway. Only clusters which contained at least 4 nodes were included in this table.

Supplementary Table S2. Number of FDA approved therapeutics predicted to target pathways. Enrichment Map post analysis was used to identify drugs from the Drug Bank database: October_1_2019_Human_DrugBank_approved_entrezgene.gmt (available at http://download.baderlab.org/EM_Genesets/October_01_2019/Human/Entrezgene/). Potential therapeutics are defined as having an overlap of at least 3 genes between an enriched geneset and drugset.

Cluster Number	Cluster Name	Number of Drug Hits
1	TLR/MYD88 Signaling	37
2	Oxidative Phosphorylation	30
3	Regulation of Adaptive Immunity	15
4	Translation	17
5	Response to Bacteria	25
6	Regulation of Apoptic Signaling	11
7	Regulation of MAPK Signaling	88
8	HIV-NEF Signaling and TNF Signaling	14
9	Phosphatidylinositol Biosynthesis	20
10	Leukocyte Chemotaxis	0
11	Cillium Organization	3
12	Regulation of Leukocyte Chemotaxis	10
13	Fatty Acid Beta Oxidation	19
14	Regulation of Blood Coagulation	7
15	Response to TNF and IL-1	12
16	Cell Death Signaling	34
17	IL-2 and GMCSF Signaling	13
18	Amino Acid Metabolism	28
19	DNA Repair	1
20	Epigenetic Regulation of Expression	1
21	TCR,BCR,IL-2 Signaling	8
22	PLC Signaling	0
23	Regulation of Protein Import to Nucleus	8
24	Regulation of FGFR Signaling	20
25	IL-12 and IL-23 Signaling	0
26	Regulation of Alternative Splicing	1
27	S1P Signaling	0
28	DNA Repair-1	1
29	DNA Repair-2	11
30	Keratinization	0
31	Regulation of ROS Metabolism/Biosynthesis	1
32	Regulation of JAK-STAT Signaling	42
33	Vesicle Transport	0
34	Regulation of Vasodilation	5
35	MAPK Signaling	0
36	Symbiont Host Modulation	0
37	unknown	11
38	Adherens Junctions	8

39	Protein Localization to Vacuole	10
40	Tubulin Folding	2
41	Response to Hormone	20
42	ERBB and TRK Receptor Signaling	5
43	Response to Inflammatory Cytokine	15
44	Cofactor Metabolic Process	0
45	TCR and BCR Signal Transduction	0
46	Leukocyte Mediated Cytotoxicity	0
47	Cholesterol Biosynthesis	0
48	Regulation of Epithelial and Endothelial Apoptosis	15
49	Amino Acid Metabolism	3
50	Intracellular Transport	6
51	Leukocyte Homeostasis	2
52	AP1 and Fra Pathway	47
53	unknown	1
54	Regulation of Endocrine Process	17
55	Thromboxane and IL-8 Signaling	0
56	NFAT and TCR Pathway	0
57	Heart Morphogenesis	0

Supplementary Data. 1

Donor lungs in the EVLP group were significantly older than lungs in the transplant dataset (Table 1). As a proxy, we analyzed differential gene expression between the 10 youngest and 10 oldest donor lungs in the transplant and evlp datasets at both pre and post timepoints (4 comparisons total) using limma (5). In the transplant dataset the mean age of young lung group was 24 (±8.26) and the mean age of old lung group was 67.18 (±5.016). In the EVLP dataset the mean age of young lung group was 19.1 (±3.84), and the mean age of old lung group was 58.55 (±4.344). No differential gene expression was found at FDR<0.05, and a few genes were found at FDR<0.10 (see below).

	Transplant	**EVLP**
Pre	**0 DGE @FDR<0.05, 0 DGE @FDR<0.10**	**0 DGE @FDR<0.05, 5 DGE @FDR<0.10**
Post	**0 DGE @FDR<0.05, 6 DGE @FDR<0.10**	**0 DGE @FDR<0.05, 0 DGE @FDR<0.10**

References

1. R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

2. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP, Beazer-barclay YD, Antonellis KJ, Speed TP. Exploration , Normalization , and Summaries of High Density Oligonucleotide Array Probe Level Data. *Biostatistics* 2018;4:249–264.

3. Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. *Bioinformatics* 2010;26:2363–2367.

4. Dai M, Wang P, Boyd AD, Kostov G, Athey B, Jones EG, Bunney WE, Myers RM, Speed TP, Akil H, Watson SJ, Meng F. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res* 2005;33:e175–e175.

5. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;43:e47.

6. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B* 1995;57:289–300.

7. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci* 2005;102:15545–15550.

8. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res* 2003;13:2498–2504.

9. Reimand J, Isserlin R, Voisin V, Kucera M, Tannus-Lopes C, Rostamianfar A, Wadi L, Meyer M, Wong J, Xu C, Merico D, Bader GD. Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat Protoc* 2019;14:482–517.

10. Kucera M, Isserlin R, Arkhangorodsky A, Bader GD. AutoAnnotate: A Cytoscape app for summarizing networks with semantic annotations. *F1000Research* 2016;5:1717.

11. January Weiner (2017). pca3d: Three Dimensional PCA Plots. R package version 0.10.

12. Kang CH, Anraku M, Cypel M, Sato M, Yeung J, Gharib SA, Pierre AF, De Perrot M, Waddell TK, Liu M, Keshavjee S. Transcriptional signatures in donor lungs from donation after cardiac death vs after brain death: A functional pathway analysis. *J Hear Lung Transplant* 2011;30:289–298.

13. Yeung JC, Zamel R, Klement W, Bai X-H, Machuca TN, Waddell TK, Liu M, Cypel M, Keshavjee S. Towards donor lung recovery-gene expression changes during ex vivo lung perfusion of human lungs. *Am J Transplant* 2018;18:1518–1526.