Technical Report

Joshua Ellis

CSC 7300

**Introduction**

Developing a new drug can take up to 15 years and cost around $2 billion dollars to bring the initial idea all the way to final approval and sale (Nosengo). This cost and time commitment means that to develop a drug it must be certain that a disease is common enough to warrant the significant investment. Many drugs never reach the final stage of development and represent a wasted investment. Some of these drugs are have been approved for use on a different disease or for use but are not effective. These drugs can be repurposed to treat lesser known disease (Ashburn). The paper “A common rejection module (CRM) for acute rejection across multiple organs identifies novel therapeutics for organ transplantation” (Khatri) applies this philosophy to innovating of drug treatments for organ rejections. In the paper they analyzed heart, lung, liver, and kidney rejection data to find overexpressed genes. They then found existing drugs (atorvastatin and dasatinib) that regulated the most overexpressed genes. The results from the paper were replicated using a few of the same methods as Khatri and in R.

**Methods**

Data sets GSE1563, GSE21374, GSE36059, GSE25092, and GSE50058 were obtained through GEO and were analyzed using GEO2R. This analysis provided the base p-values from each experiment. The effect size of each differentially expressed gene from each experiment was used to determine if the samples were comparable (Leeds.ac.uk.). This was done using the effsize package in R and the fold change values from GEO2R analysis. The resulting p-values were combined using Fisher’s method, also known as sum of logs. Fisher’s method (Udel) takes the logarithm of each p value and adds them over k experiments. In this case k was 5. And these new p-values were then ranked to determine the overexpressed genes in the renal rejection patients (Elston).

**Data description**

Five data sets from five experiments [Flechner,Einecke,Reeve,Naesens,Khatri]. All were taken from Kidneys and split between Acute Rejection and Stable patients[Khatri].

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **GEO ID** | **GEO Platform ID** | **Microarray** | **Samples (AR/STA)** | **Samples (Total)** |
| GSE1563 | GPL96 | Affymetrix HG U133A | 7/9 | 16 |
| GSE21374 | GPL570 | Affymetrix HG U133 Plus 2.0 | 76/206 | 282 |
| GSE36059 | GPL570 | Affymetrix HG U133 Plus 2.0 | 122/221 | 348 |
| GSE25092 | GPL570 | Affymetrix HG U133 Plus 2.0 | 24/96 | 120 |
| GSE50058 | GPL570 | Affymetrix HG U133 Plus 2.0 | 43/58 | 101 |
| **Total** | **2** | **-** | **272/664** | **936** |

**Results**

**Significant Genes**

The results of the analysis showed 43 overexpressed genes compared to the 10 overexpressed genes in the renal samples in the paper. With 7 genes (ISG20,CD8,CXCL10,CXCL9,NKG7,PSMB9, and TAP2) showing up in both analyses. Depending on the number of genes chosen from each study (500-5000 from each set) the number of significant genes varies from 43 to 434 genes. Included in the Appendix is a top 100 list from the analysis that included the maximum genes (434). This was not the method used in the paper but shows the variability in analyzing this data.

**Pathway Analysis**

Pathway analysis was conducted using reactome. As the figure shows the significant pathways include functions relating to the immune system, the cell cycle, DNA replication and repair, and apoptosis.

**Discussion**

This difference illustrates the differences between the paper’s analysis and the analysis preformed exclusively in R. The meta-analyses of both effect size and p-value can take different parameters and limits to calculate significance could have contributed to this difference. When different number of genes are chosen from each data set, as is necessary to analyze all the data and remove noise, it has a large effect on the number of genes found but the list that overlaps stays fairly consistent. With between 5 and 7 similar genes found in each paper. The pathway analysis shows that the genes found are related to the immune system and the life cycle of the cell. This is consistent for the graft rejection. It would be expected that genes related to immune reject of foreign objects would be expressed. And these cells dying would be expressed in the function of DNA and apoptosis.

**Bibliography**

Elston, R. (1991). On Fisher's Method of Combining p-Values. *Journal of Biomedical*, [online] 33, pp.339-345. Available at: http://darwin.cwru.edu/ref/view.php?id=316&article=Elston+Reprints.

Khatri, P., Roedder, S., Kimura, N., De Vusser, K., Morgan, A., Gong, Y., Fischbein, M., Robbins, R., Naesens, M., Butte, A. and Sarwal, M. (2013). A common rejection module (CRM) for acute rejection across multiple organs identifies novel therapeutics for organ transplantation. *The Journal of Experimental Medicine*, 210(11), pp.2205-2221.

Leeds.ac.uk. (2017). *It's the effect size, stupid: what effect size is and why it is important*. [online] Available at: https://www.leeds.ac.uk/educol/documents/00002182.htm [Accessed 23 Aug. 2017].

meta-analysis.com. (2017). *Effect Sizes Based on Means*. [online] Available at: https://www.meta-analysis.com/downloads/Meta-analysis%20Effect%20sizes%20based%20on%20means.pdf [Accessed 19 Jul. 2017].

Nosengo, N. (2017). *Can you teach old drugs new tricks?*. [online] nature.com. Available at: https://www.nature.com/news/can-you-teach-old-drugs-new-tricks-1.20091 [Accessed 2 Jul. 2017].

rejection, T. (2017). *Transplant rejection: MedlinePlus Medical Encyclopedia*. [online] Medlineplus.gov. Available at: https://medlineplus.gov/ency/article/000815.htm [Accessed 14 Aug. 2017].

Sharma, V. (2012). Drug Repositioning: A Faster Path to Drug Discovery. *Advances in Pharmacoepidemiology & Drug Safety*, 01(06).

Udel.edu. (2017). *Handbook of Biological Statistics: Fisher's exact test of independence*. [online] Available at: http://udel.edu/~mcdonald/statfishers.html [Accessed 2 Aug. 2017].

**APPENDIX**

|  |  |
| --- | --- |
| 1.71E-57 | FAM102A |
| 1.64E-18 | ALB |
| 1.66E-19 | ATP6V1H |
| 1.47E-19 | C11orf95 |
| 1.11E-19 | AURKAIP1 |
| 1.09E-19 | ATP5S |
| 1.55E-20 | ASPA |
| 1.44E-20 | ATP6V0A4 |
| 1.06E-20 | ATP6V1G2-DDX39B///SNORD84///DDX39B |
| 1.16E-21 | APLP2 |
| 1.12E-21 | ARHGEF17 |
| 1.30E-25 | ADA |
| 1.54E-26 | CARD16 |
| 1.61E-28 | ALOX5 |
| 1.31E-29 | C2orf40 |
| 1.33E-30 | ARHGAP9 |
| 1.34E-34 | IGK///IGKC |
| 1.67E-35 | SCD5 |
| 1.50E-35 | SRGN |
| 1.50E-35 | NEBL |
| 1.26E-35 | POLR1B |
| 1.22E-35 | MAN2B2 |
| 1.22E-35 | MTHFD2 |
| 1.09E-35 | NCF2 |
| 1.47E-36 | MED31 |
| 1.22E-36 | ADAMDEC1 |
| 1.20E-36 | MAP2K7 |
| 1.06E-36 | MLX |
| 1.71E-37 | SPECC1L |
| 1.55E-37 | LYZ |
| 1.22E-37 | MCUB |
| 1.08E-37 | NRIP1 |
| 1.48E-38 | MS4A7 |
| 1.46E-38 | RIOK3 |
| 1.26E-38 | PTMS |
| 1.16E-38 | PTTG1 |
| 1.01E-38 | RAP2C |
| 1.32E-39 | SNHG16///SNORD1C///SNORD1A |
| 1.19E-40 | PLAC8 |
| 1.39E-41 | SLC5A3 |
| 1.31E-41 | 1-Mar |
| 1.05E-41 | SLC12A3 |
| 1.49E-42 | INSR |
| 1.00E-43 | PSMB9 |
| 1.53E-44 | LOC729966///PDE4C |
| 1.31E-44 | KMO |
| 1.19E-44 | PLEK |
| 1.13E-44 | RAC2 |
| 1.41E-45 | LY9 |
| 1.42E-46 | LINC01187 |
| 1.19E-46 | CD3D |
| 1.19E-46 | CCL5 |
| 1.17E-46 | TACC1 |
| 1.43E-47 | UBE3C |
| 1.36E-47 | TLN2 |
| 1.10E-47 | SLC25A23 |
| 1.33E-48 | TMEM198B |
| 1.32E-48 | TEF |
| 1.26E-48 | TSPAN12 |
| 1.06E-48 | TNFAIP3 |
| 1.64E-49 | RER1 |
| 1.30E-49 | CPM |
| 1.52E-50 | TRG-AS1 |
| 1.07E-51 | TUB |
| 1.63E-52 | FABP1 |
| 1.48E-53 | TYROBP |
| 1.38E-54 | KLRB1 |
| 1.24E-54 | HRG |
| 1.14E-54 | ENTPD5 |
| 1.26E-55 | FAM210A |
| 1.18E-55 | IFNAR2 |
| 1.68E-56 | EFHD1 |
| 1.23E-56 | FCER1G |
| 1.59E-57 | FAM162A |
| 1.12E-57 | TAP1 |
| 1.02E-57 | EXOC3 |
| 1.45E-58 | FAAH |
| 1.43E-58 | EVI2B |
| 1.12E-58 | GRB2 |
| 1.70E-59 | HDAC6 |
| 1.54E-60 | CLC |
| 1.28E-60 | TLR8 |
| 1.17E-60 | CTSS |
| 1.54E-61 | FCRL5 |
| 1.35E-61 | CXCL6 |
| 1.15E-61 | CD84 |
| 1.09E-61 | CELSR1 |
| 1.03E-62 | GPATCH4 |
| 1.37E-63 | EVI2A |
| 1.24E-63 | CRTAM |
| 1.69E-65 | CLEC7A |
| 1.40E-65 | IRF1 |
| 1.53E-66 | GPR65 |
| 1.23E-66 | CD86 |
| 1.08E-68 | TRBC1 |
| 1.20E-71 | FYB |
| 1.68E-77 | CXCL10 |
| 1.17E-83 | GBP2 |
| 1.41E-94 | CXCL11 |