## **1.Materials and Methods**

## 1.1 Differentially Expresses Genes

With gene expression result of the GSE32719 data set, we detect Differentially Expressed Gene (**DEG**) between conditions from young-old and young-middle aged samples. We find the up-regulated and down-regulated differentially expressed genes between conditions.

We also investigate up and down regulated gene expressions between **young-old, young-middle**, and **middle-old** conditions.

We detected 17 differentially expressed genes of down regulation, whereas we find 19 up-regulated genes between high and low conditions.

|  |  |  |
| --- | --- | --- |
| Conditions Compared | down-regulated | up-regulated |
| young-old | 9 | 21 |
| young-middle | 15 | 51 |
| middle-old | - | - |

**Table:** Summary of differentially expressed genes by conditions.

## 1.2 Gene Ontology Enrichment Analysis

After detecting the differentially over (up) and under (down) expressed genes between conditions including high and low groups**, GOrilla** is used to identify and visualize enriched GO terms in ranked lists of genes (Eden et al, 2009 & Eden et al 2007). We do the enrichment analyses for genes in the intersection of all up and down-regulated groups. We further analyze up and down- regulated genes between high and low conditions. We have used approximately 21,000 genes as a background gene reference with two-rank listed analyses in GOrilla.

**Gorilla has calculated the 'P-value'** which is the enrichment p-value computed according to the mHG or HG model. This p-value is not corrected for multiple testing of 14548 GO terms.  
  
**The calculated q-value is called 'FDR q-value'** which is the correction of the above p-value for multiple testing using the Benjamini and Hochberg (1995) method.

For the ith term (ranked according to p-value) the FDR q-value is,

**The value for Enrichment (N, B, n, b)** is defined as follows:

N - is the total number of genes  
B - is the total number of genes associated with a specific GO term  
n - is the number of genes in the top of the user's input list or in the target set when appropriate  
b - is the number of genes in the intersection

**Genes:** For each GO term you can see the list of associated genes that appear in the optimal top of the list.  
Each gene name is specified by gene symbol followed by a short description of the gene

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **GO Term** | **Description** | **P-value** | **FDR q-value** | **Enrichment** | **N** | **B** | **n** | **b** | **Genes** |
| GO:0048545 | response to steroid hormone | 5.51E-05 | 2.92E-01 | 16.3 | 880 | 12 | 18 | 4 | LOX , CASP9 , FLT3 , SPP1 |
| GO:0014070 | response to organic cyclic compound | 6.60E-05 | 1.75E-01 | 10.19 | 880 | 24 | 18 | 5 | LOX , CASP9 , ADA, FLT3 , SPP1 |
| GO:1901361 | organic cyclic compound catabolic process | 3.15E-04 | 5.57E-01 | 10.86 | 880 | 18 | 18 | 4 | AGO3 , FAH , ADA , SPP1 |
| GO:1901575 | organic substance catabolic process | 3.43E-04 | 4.54E-01 | 5.64 | 880 | 52 | 18 | 6 | AGO3 , FUT2 , FAH , ADA , GLA , SPP1 |
| GO:0071407 | cellular response to organic cyclic compound | 7.90E-04 | 8.39E-01 | 14.67 | 880 | 10 | 18 | 3 | CASP9 ,FLT3 , SPP1 |
| GO:0044282 | small molecule catabolic process | 7.90E-04 | 6.99E-01 | 14.67 | 880 | 10 | 18 | 3 | FUT2 , FAH , ADA |
| GO:0042127 | regulation of cell proliferation | 9.15E-04 | 6.93E-01 | 4.73 | 880 | 62 | 18 | 6 | AGO3 , FUT2 , HPGDS , ADA , FLT3 , NMB |
| GO:0009056 | catabolic process | 9.15E-04 | 6.07E-01 | 4.73 | 880 | 62 | 18 | 6 | AGO3 , FUT2, FAH , ADA , GLA , SPP1 |

|  |
| --- |
| Table 1: Summary of GO pathways of all up-regulated genes (in total 21 genes) from the conditions between young and old aged samples. |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **GO Term** | **Description** | **P-value** | **FDR q-value** | **Enrichment** | **N** | **B** | **n** | **b** | **Genes** |
| GO:0032496 | response to lipopolysaccharide | 8.86E-05 | 4.60E-01 | 29.03 | 871 | 10 | 9 | 3 | FOS, KMO, SELP |
| GO:0002237 | response to molecule of bacterial origin | 8.86E-05 | 2.30E-01 | 29.03 | 871 | 10 | 9 | 3 | FOS, KMO, SELP |
| GO:0034654 | nucleobase-containing compound biosynthetic process | 5.95E-04 | 1.00E+00 | 8.8 | 871 | 44 | 9 | 4 | FOS, NPR2, TOPORS, KMO |
| GO:0018130 | heterocycle biosynthetic process | 6.50E-04 | 8.45E-01 | 8.6 | 871 | 45 | 9 | 4 | FOS, NPR2, TOPORS, KMO |
| GO:0019438 | aromatic compound biosynthetic process | 6.50E-04 | 6.76E-01 | 8.6 | 871 | 45 | 9 | 4 | FOS, NPR2, TOPORS, KMO |
| GO:0044271 | cellular nitrogen compound biosynthetic process | 7.09E-04 | 6.14E-01 | 8.42 | 871 | 46 | 9 | 4 | FOS, NPR2, TOPORS, KMO |
| GO:1901362 | organic cyclic compound biosynthetic process | 8.37E-04 | 6.22E-01 | 8.06 | 871 | 48 | 9 | 4 | FOS, NPR2, TOPORS, KMO |

Table 2: Summary of GO pathways of all down-regulated genes (in total 9 genes) from the conditions between young and old aged samples.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **GO Term** | **Description** | **P-value** | **FDR q-value** | **Enrichment** | **N** | **B** | **n** | **b** | **Genes** |
| GO:0006413 | translational initiation | 8.10E-05 | 4.38E-01 | 22.55 | 902 | 3 | 40 | 3 | EIF3C, EIF3CL, EIF3G |
| GO:0001755 | neural crest cell migration | 8.10E-05 | 2.19E-01 | 22.55 | 902 | 3 | 40 | 3 | SEMA4A, PITX2, SHH |
| GO:0001732 | formation of cytoplasmic translation initiation complex | 8.10E-05 | 1.46E-01 | 22.55 | 902 | 3 | 40 | 3 | EIF3C, EIF3CL, EIF3G |
| GO:0022618 | ribonucleoprotein complex assembly | 2.05E-04 | 2.77E-01 | 11.28 | 902 | 8 | 40 | 4 | EIF3C, EIF3CL, SRPK3, EIF3G |
| GO:0071826 | ribonucleoprotein complex subunit organization | 3.57E-04 | 3.86E-01 | 10.02 | 902 | 9 | 40 | 4 | EIF3C, EIF3CL, SRPK3, EIF3G |
| GO:0034622 | cellular protein-containing complex assembly | 4.92E-04 | 4.43E-01 | 5.41 | 902 | 25 | 40 | 6 | EIF3C, NDUFB11, EIF3CL, DIAPH1, SRPK3, EIF3G |
| GO:0090185 | negative regulation of kidney development | 7.61E-04 | 5.87E-01 | 13.53 | 902 | 5 | 40 | 3 | SHH, STAT1, TACSTD2 |
| GO:0071407 | cellular response to organic cyclic compound | 8.76E-04 | 5.92E-01 | 8.2 | 902 | 11 | 40 | 4 | RALB, EL1, DIAPH1, STAT1 |

Table 3: Summary of GO pathways of all up-regulated genes (in total 51 genes) from the conditions between young and middle aged samples.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **GO Term** | **Description** | **P-value** | **FDR q-value** | **Enrichment** | **N** | **B** | **n** | **b** |
| GO:0051172 | negative regulation of nitrogen compound metabolic process | 7.79E-06 | 4.07E-02 | 5.98 | 875 | 90 | 13 | 8 |
| GO:2000113 | negative regulation of cellular macromolecule biosynthetic process | 9.85E-06 | 2.57E-02 | 9.85 | 875 | 41 | 13 | 6 |
| GO:0010605 | negative regulation of macromolecule metabolic process | 1.19E-05 | 2.07E-02 | 5.67 | 875 | 95 | 13 | 8 |
| GO:0031324 | negative regulation of cellular metabolic process | 1.19E-05 | 1.55E-02 | 5.67 | 875 | 95 | 13 | 8 |
| GO:0071495 | cellular response to endogenous stimulus | 1.39E-05 | 1.45E-02 | 13.46 | 875 | 25 | 13 | 5 |
| GO:0010558 | negative regulation of macromolecule biosynthetic process | 1.51E-05 | 1.32E-02 | 9.18 | 875 | 44 | 13 | 6 |
| GO:0051253 | negative regulation of RNA metabolic process | 1.51E-05 | 1.13E-02 | 9.18 | 875 | 44 | 13 | 6 |
| GO:0071310 | cellular response to organic substance | 1.73E-05 | 1.13E-02 | 8.97 | 875 | 45 | 13 | 6 |
| GO:0031327 | negative regulation of cellular biosynthetic process | 1.98E-05 | 1.15E-02 | 8.78 | 875 | 46 | 13 | 6 |
| GO:0009892 | negative regulation of metabolic process | 2.22E-05 | 1.16E-02 | 5.23 | 875 | 103 | 13 | 8 |
| GO:0009890 | negative regulation of biosynthetic process | 2.56E-05 | 1.21E-02 | 8.41 | 875 | 48 | 13 | 6 |
| GO:0010033 | response to organic substance | 2.67E-05 | 1.16E-02 | 6.37 | 875 | 74 | 13 | 7 |
| GO:0000122 | negative regulation of transcription by RNA polymerase II | 3.57E-05 | 1.44E-02 | 11.22 | 875 | 30 | 13 | 5 |
| GO:0010629 | negative regulation of gene expression | 3.67E-05 | 1.37E-02 | 7.92 | 875 | 51 | 13 | 6 |
| GO:0045934 | negative regulation of nucleobase-containing compound metabolic process | 3.67E-05 | 1.28E-02 | 7.92 | 875 | 51 | 13 | 6 |
| GO:0010628 | positive regulation of gene expression | 4.16E-05 | 1.36E-02 | 5.96 | 875 | 79 | 13 | 7 |
| GO:0070887 | cellular response to chemical stimulus | 4.61E-05 | 1.42E-02 | 7.62 | 875 | 53 | 13 | 6 |
| GO:0042221 | response to chemical | 5.81E-05 | 1.68E-02 | 5.68 | 875 | 83 | 13 | 7 |
| GO:2000112 | regulation of cellular macromolecule biosynthetic process | 6.16E-05 | 1.69E-02 | 3.83 | 875 | 158 | 13 | 9 |
| GO:0010556 | regulation of macromolecule biosynthetic process | 7.22E-05 | 1.88E-02 | 3.76 | 875 | 161 | 13 | 9 |
| GO:0010604 | positive regulation of macromolecule metabolic process | 1.03E-04 | 2.56E-02 | 4.27 | 875 | 126 | 13 | 8 |
| GO:0045892 | negative regulation of transcription, DNA-templated | 1.03E-04 | 2.46E-02 | 9.1 | 875 | 37 | 13 | 5 |
| GO:0031326 | regulation of cellular biosynthetic process | 1.14E-04 | 2.59E-02 | 3.56 | 875 | 170 | 13 | 9 |
| GO:0009719 | response to endogenous stimulus | 1.18E-04 | 2.57E-02 | 8.86 | 875 | 38 | 13 | 5 |
| GO:0009889 | regulation of biosynthetic process | 1.52E-04 | 3.18E-02 | 3.44 | 875 | 176 | 13 | 9 |
| GO:1902679 | negative regulation of RNA biosynthetic process | 1.53E-04 | 3.06E-02 | 8.41 | 875 | 40 | 13 | 5 |
| GO:1903507 | negative regulation of nucleic acid-templated transcription | 1.53E-04 | 2.95E-02 | 8.41 | 875 | 40 | 13 | 5 |
| GO:0010468 | regulation of gene expression | 1.60E-04 | 2.98E-02 | 3.42 | 875 | 177 | 13 | 9 |
| GO:0048511 | rhythmic process | 2.05E-04 | 3.69E-02 | 22.44 | 875 | 9 | 13 | 3 |
| GO:0010557 | positive regulation of macromolecule biosynthetic process | 2.14E-04 | 3.73E-02 | 5.85 | 875 | 69 | 13 | 6 |
| GO:0014070 | response to organic cyclic compound | 2.23E-04 | 3.75E-02 | 11.71 | 875 | 23 | 13 | 4 |
| GO:0009893 | positive regulation of metabolic process | 2.77E-04 | 4.52E-02 | 3.74 | 875 | 144 | 13 | 8 |
| GO:0071407 | cellular response to organic cyclic compound | 2.90E-04 | 4.59E-02 | 20.19 | 875 | 10 | 13 | 3 |
| GO:0051716 | cellular response to stimulus | 2.93E-04 | 4.50E-02 | 4.44 | 875 | 106 | 13 | 7 |
| GO:0031328 | positive regulation of cellular biosynthetic process | 3.45E-04 | 5.14E-02 | 5.38 | 875 | 75 | 13 | 6 |
| GO:0045935 | positive regulation of nucleobase-containing compound metabolic process | 4.00E-04 | 5.80E-02 | 5.24 | 875 | 77 | 13 | 6 |
| GO:0009891 | positive regulation of biosynthetic process | 4.00E-04 | 5.64E-02 | 5.24 | 875 | 77 | 13 | 6 |
| GO:0051252 | regulation of RNA metabolic process | 5.46E-04 | 7.49E-02 | 3.41 | 875 | 158 | 13 | 8 |
| GO:0032922 | circadian regulation of gene expression | 6.07E-04 | 8.12E-02 | 44.87 | 875 | 3 | 13 | 2 |
| GO:0045444 | fat cell differentiation | 6.07E-04 | 7.92E-02 | 44.87 | 875 | 3 | 13 | 2 |
| GO:0007584 | response to nutrient | 6.07E-04 | 7.72E-02 | 44.87 | 875 | 3 | 13 | 2 |
| GO:0051173 | positive regulation of nitrogen compound metabolic process | 7.23E-04 | 8.99E-02 | 3.86 | 875 | 122 | 13 | 7 |
| GO:0033993 | response to lipid | 8.39E-04 | 1.02E-01 | 8.41 | 875 | 32 | 13 | 4 |

Table 4: Summary of GO pathways of all down-regulated genes (in total 15 genes) from the conditions between young and middle aged samples.

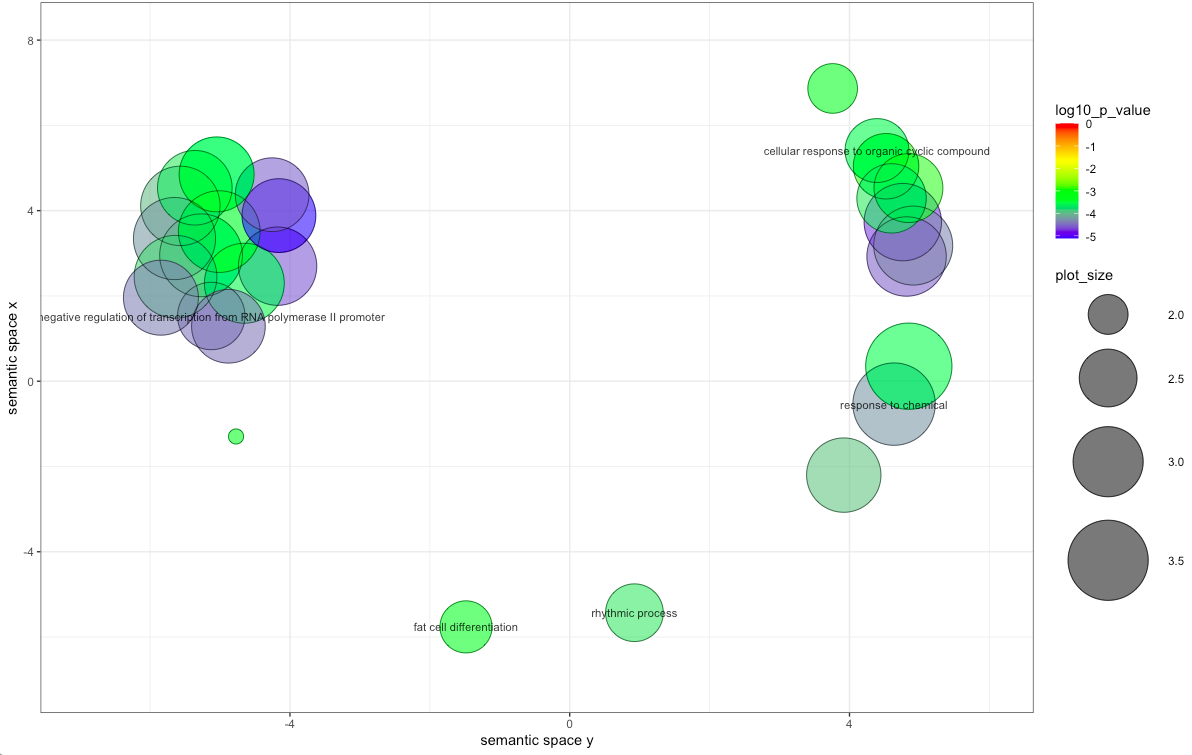
****

Figure 1: Plot of GO pathways of significant biological processes of all down-regulated genes (in total 15 genes) from the conditions between young and middle aged samples. Negative regulation of transcription from RNA polymerase II promoter results as the most significant biological process with a log10(P-value) < -4 .

**REFERENCE:**

Tarazona, Sonia, et al. "Differential expression in RNA-seq: a matter of depth." *Genome research* 21.12 (2011): 2213-2223.

Eran Eden\*, Roy Navon\*, Israel Steinfeld, Doron Lipson and Zohar Yakhini. *"GOrilla: A Tool For Discovery And Visualization of Enriched GO Terms in Ranked Gene Lists"*, BMC Bioinformatics 2009, 10:48.

 Eran Eden, Doron Lipson, Sivan Yogev, Zohar Yakhini. "**Discovering Motifs in Ranked Lists of DNA sequence**s", PLoS Computational Biology, 3(3):e39, 2007.

[1] M. Fraga and M. Esteller, “Fraga, M. F. & Esteller, M. Epigenetics and aging: the targets and the marks. Trends Genet. 23, 413-418,” *Trends Genet.*, vol. 23, pp. 413–418, Sep. 2007, doi: 10.1016/j.tig.2007.05.008.

[2] M. P. Rodero *et al.*, “Immune surveillance of the lung by migrating tissue monocytes,” *Elife*, vol. 4, pp. e07847–e07847, Jul. 2015, doi: 10.7554/eLife.07847.

[3] A. Karadağ, B. Altinok, T. Özkan, ve Y. Hekmatshoar, “Kemik İliği Stroması: Hücreleri ve Mikroçevresi”, s. 12.

[4] J. M. Perry ve L. Li, “Disrupting the Stem Cell Niche: Good Seeds in Bad Soil”, *Cell*, c. 129, sy 6, ss. 1045-1047, Haz. 2007.

[5] E. Karaöz, “Ovalı E”, *Kök Hücreler Derya Kitabevi Trabzon*, 2004.

[6] J. E. Till ve E. A. McCulloch, “A direct measurement of the radiation sensitivity of normal mouse bone marrow cells”, *Radiat. Res.*, c. 14, sy 2, ss. 213-222, 1961.

[7] A. U. Ural, “Hematopoetik Kök Hücre”, *Turk. Klin. J. Surg. Med. Sci.*, c. 2, sy 43, ss. 5-10, 2006.

[8] G. Aydın, “Değişik Kaynaklı MKH’lerin Detaylı Karakterizasyonu, Lenfohematopoezi Destekleyici Özellikleri ve İzolasyonunda CD271 Antijeninin Önemi”, 2018.

[9] G. Gürhan, “Deneysel Balb/C Ova Akut Astim Modelinde Fare Kemik İliği Kaynakli Mezenkimal Kök Hücrelerin Farkli Uygulama Yollarinin Astim Modeli Üzerine Etkisinin Araştirilmasi”.

[10] A. S. Çakmak, “Biyofiziksel ve Biyokimyasal Uyaranlarla Desteklenmiş Doku İskeleleri ile Mezenkimal Kök Hücrelerin Osteojenik Farklılaşmasının İncelenmesi”, 2014.

[11] D. J. Rossi *vd.*, “Cell intrinsic alterations underlie hematopoietic stem cell aging”, *Proc. Natl. Acad. Sci.*, c. 102, sy 26, ss. 9194-9199, Haz. 2005.

[12] W. W. Pang *et al.*, “Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 108, no. 50, pp. 20012–20017, 2011, doi: 10.1073/pnas.1116110108.

[13] S. Davis and P. Meltzer, “GEOquery: A bridge between the Gene Expression Omnibus (GEO) and BioConductor,” *Bioinformatics*, vol. 23, pp. 1846–1847, Aug. 2007, doi: 10.1093/bioinformatics/btm254.

[14] R. Ihaka and R. Gentleman, “R: A Language for Data Analysis and Graphics,” *J. Comput. Graph. Stat.*, vol. 5, no. 3, pp. 299–314, Sep. 1996, doi: 10.1080/10618600.1996.10474713.

[15] R Core Team, “R: A language and environment for statistical computing and Foundation for Statistical Computing.” 2020.

[16] G. R. Warnes *et al.*, “gplots: Various R programming tools for plotting data,” *R Packag. version*, vol. 2, no. 4, p. 1, 2009.

[17] A. D. Yates *et al.*, “Ensembl 2020,” *Nucleic Acids Res.*, vol. 48, no. D1, pp. D682–D688, Nov. 2019, doi: 10.1093/nar/gkz966.

[18] E. Eden, R. Navon, I. Steinfeld, D. Lipson, and Z. Yakhini, “GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists,” *BMC Bioinformatics*, vol. 10, no. 1, pp. 1–7, 2009.

[19] D. W. Huang *et al.*, “DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists,” *Nucleic Acids Res.*, vol. 35, no. suppl\_2, pp. W169–W175, 2007.

[20] B. T. Sherman and R. A. Lempicki, “Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources,” *Nat. Protoc.*, vol. 4, no. 1, p. 44, 2009.