

فاز دوم پروژه بیوانفورماتیک

نگار عسکری ۹۹۱۰۵۶۱۲

سوگند صالحی ۹۹۱۰۵۵۵۹

۱. کد کامل با کامنت ها پیوست شده است. توضیحات اضافه تر:

ابتدا اسم گروه‌ها که به شکل لیستی از رشته‌ها بوده را تبدیل به فاکتور می‌کنیم. فاکتور ها عملا معیاری هستند که در ظاهر همان رشته به نظر می‌آیند اما در باطن عدد می‌باشند و می‌توان برای آنها تابع levels را صدا کرد.

سپس به ساخت ماتریس دیزاین می‌پردازیم که عملا در هر سطر آن یک سمپل وجود دارد و ستون‌های آن اسمشان همان گروه‌های ما هستند و در سطور با ۰ و ۱ مشخص می‌شود که این سمپل عضو کدام گروه می‌باشد.

سپس یک مدل خطی را در نظر گرفته و داده‌ها و ماتریس دیزاین را به عنوان ورودی به آن می‌دهیم و فیت می‌کنیم.

سپس مشخص می‌کنیم که چه گروه‌هایی باید با یکدیگر مقایسه شوند (با توجه به نتایج فاز یک، ما AML و Monocytes را انتخاب کردیم) و بعد ماتریس کانتراست را می‌سازیم. سپس این کانتراست را بر روی مدل خطی فیت می‌کنیم. و نهایتا نتیجه بدست آمده را در مدل بیزین فیت می‌کنیم. مدل بیزین یک پارامتر ورودی دارد که عملا نشان می‌دهد معمولا در این تحقیقات چند درصد ژن‌ها تفاوت بیان مهمی دارند (prior). بعد این مدل با قانون بیز بر روی دیتای خودمان چک می‌کند که چه چیزی محتمل تر است و از آنجا P-Value و Adjusted P-Value محاسبه می‌شود. همچنین پارامتر fdr در adjust همان روش false discovery rate برای adjust کردن P-value می‌باشد.

در نهایت یک جدول از ژن‌هایی که بیان آن‌ها به صورت معناداری در این دو گروه متفاوت است می‌سازیم و نام ژن‌هایی که در هر گروه بیان بیشتری داشته‌اند را در فایل‌های aml_down_genes.txt و aml_up_genes.txt ذخیره می‌کنیم.

۲. از سایت Enrichr می توانیم داده هایمان را بررسی کنیم. برای AML Up Gene داریم:

Pathways:

WikiPathway 2021 Human

Retinoblastoma gene in cancer WP2446

Cell cycle WP179

G1 to S cell cycle control WP45

Histone Modifications WP2369

DNA IR-damage and cellular response via ATR WP4016

DNA Replication WP466

Effect of progerin on genes involved in Hutchinson-Gilford progeria syndrome WP4320

miRNA regulation of DNA damage response WP1530

DNA IR-double strand breaks and cellular response via ATM WP3959

DNA damage response WP707

Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	Retinoblastoma gene in cancer WP2446	1.366e-20	4.793e-18	15.38	703.57
2	G1 to S cell cycle control WP45	2.461e-13	2.880e-11	13.18	382.53
3	Cell cycle WP179	1.023e-13	1.796e-11	8.48	253.78
4	Histone Modifications WP2369	8.082e-11	7.092e-9	10.52	244.57
5	DNA Replication WP466	2.636e-8	0.000001542	11.81	206.16
6	DNA IR-damage and cellular response via ATR WP4016	1.510e-10	1.060e-8	9.06	204.93
7	Effect of progerin on genes involved in Hutchinson-Gilford progeria syndrome WP4320	9.871e-7	0.00004702	10.67	147.50
8	Ketogenesis and Ketolysis WP4742	0.001299	0.01983	19.73	131.13
9	Synthesis and Degradation of Ketone Bodies WP311	0.008269	0.1001	21.89	104.96
10	miRNA regulation of DNA damage response WP1530	0.000001072	0.00004702	6.77	93.10

Ontologies:

Human Phenotype Ontology

Myelodysplasia (HP:0002863)

Abnormality of the duodenum (HP:0002246)

Duodenal stenosis (HP:0100867)

Small intestinal stenosis (HP:0012848)

Sloping forehead (HP:0000340)

Aganglionic megacolon (HP:0002251)

Acute myeloid leukemia (HP:0004808)

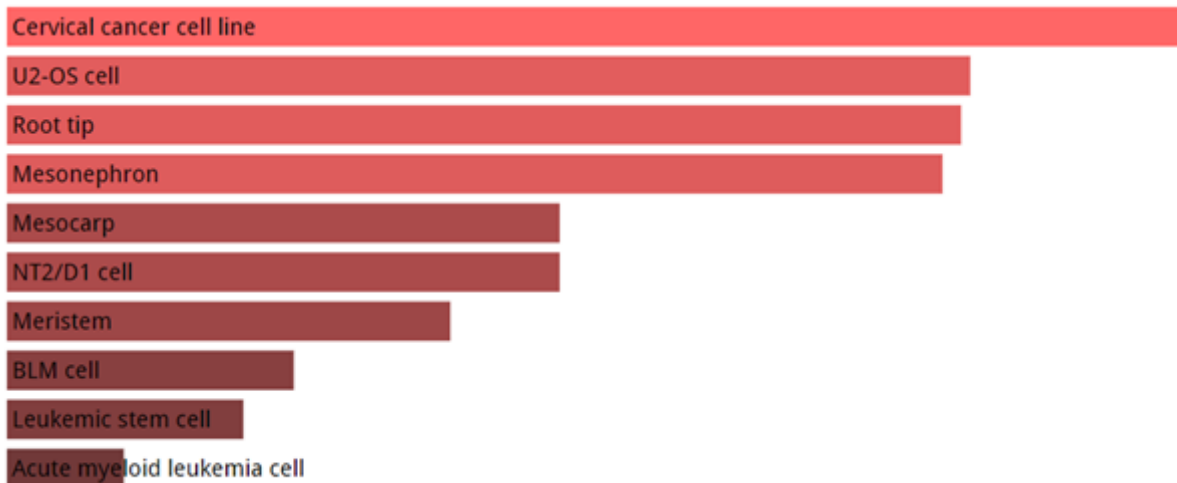
Chromosomal breakage induced by crosslinking agents (HP:0003221)

Radial deviation of finger (HP:0009466)

Clinodactyly of the 5th finger (HP:0004209)

Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	Chromosomal breakage induced by crosslinking agents (HP:0003221)	0.00004870	0.005680	32.94	327.11
2	Chromosome breakage (HP:0040012)	0.00008560	0.006194	26.35	246.81
3	Acute lymphatic leukemia (HP:0006721)	0.0001393	0.007977	21.96	194.97
4	Myelodysplasia (HP:0002863)	7.378e-8	0.00006884	10.46	171.79
5	Duodenal stenosis (HP:0100867)	0.000002459	0.0005736	11.53	148.86
6	Small intestinal stenosis (HP:0012848)	0.000002459	0.0005736	11.53	148.86
7	Abnormality of the duodenum (HP:0002246)	0.000001260	0.0005736	10.30	139.89
8	Abnormality of the heme biosynthetic pathway (HP:0010472)	0.0001037	0.006910	13.74	126.09
9	Acute myeloid leukemia (HP:0004808)	0.00004346	0.005680	11.66	117.09
10	Aplasia/Hypoplasia of the uvula (HP:0010293)	0.0001855	0.008653	11.78	101.21

Jensen TISSUES



Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	Cervical cancer cell line	4.683e-11	1.721e-8	8.31	197.64
2	U2-OS cell	4.506e-7	0.00005520	11.95	174.59
3	Root tip	2.887e-8	0.000005304	10.00	173.61
4	Mesonephron	0.00001293	0.0007308	15.25	171.65
5	Mesocarp	0.001299	0.03537	19.73	131.13
6	NT2/D1 cell	0.001299	0.03537	19.73	131.13
7	Meristem	0.000001153	0.0001211	8.74	119.54
8	BLM cell	0.001906	0.04670	16.44	102.97
9	Leukemic stem cell	0.0006041	0.02018	13.17	97.63
10	Acute myeloid leukemia cell	0.000005989	0.0004002	7.07	84.97

در موارد گزارش شده به سرطانی بودن سلول‌ها و حتی به صورت مستقیم به AML اشاره شده است.

Pathways:

Reactome 2022

Innate Immune System R-HSA-168249
Immune System R-HSA-168256
Neutrophil Degranulation R-HSA-6798695
Immunoregulatory Interactions Between A Lymphoid And A non-Lymphoid Cell R-HSA-198933
Adaptive Immune System R-HSA-1280218
Toll-like Receptor Cascades R-HSA-168898
Interleukin-10 Signaling R-HSA-6783783
CDC42 GTPase Cycle R-HSA-9013148
Toll Like Receptor 7/8 (TLR7/8) Cascade R-HSA-168181
Diseases Associated With TLR Signaling Cascade R-HSA-5602358

Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	Neutrophil Degranulation R-HSA-6798695	1.125e-24	4.030e-22	5.09	280.68
2	Innate Immune System R-HSA-168249	6.765e-29	7.272e-26	3.83	248.62
3	Immunoregulatory Interactions Between A Lymphoid And A non-Lymphoid Cell R-HSA-198933	2.430e-12	6.531e-10	6.90	184.44
4	Immune System R-HSA-168256	1.969e-27	1.058e-24	2.95	181.36
5	Formyl Peptide Receptors Bind Formyl Peptides And Many Other Ligands R-HSA-444473	0.002401	0.06144	15.81	95.36
6	Interleukin-10 Signaling R-HSA-6783783	0.000004150	0.0006373	7.59	94.05
7	MyD88 Deficiency (TLR2/4) R-HSA-5602498	0.0002813	0.02198	11.00	89.98
8	Advanced Glycosylation Endproduct Receptor Signaling R-HSA-879415	0.0009869	0.03536	11.72	81.14
9	Trafficking And Processing Of Endosomal TLR R-HSA-1679131	0.0009869	0.03536	11.72	81.14
10	IRAK4 Deficiency (TLR2/4) R-HSA-5603041	0.0003778	0.02256	10.16	80.06

Ontologies:

GO Biological Process 2021

neutrophil activation involved in immune response (GO:0002283)

neutrophil degranulation (GO:0043312)

neutrophil mediated immunity (GO:0002446)

positive regulation of cytokine production (GO:0001819)

inflammatory response (GO:0006954)

innate immune response (GO:0045087)

immune response-activating cell surface receptor signaling pathway (GO:0002429)

regulation of interleukin-6 production (GO:0032675)

response to lipopolysaccharide (GO:0032496)

cellular response to molecule of bacterial origin (GO:0071219)

Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	complement receptor mediated signaling pathway (GO:0002430)	9.446e-7	0.0001441	31.75	440.44
2	immune response-activating cell surface receptor signaling pathway (GO:0002429)	5.130e-9	0.000002347	18.99	362.55
3	negative regulation of dendritic cell differentiation (GO:2001199)	0.0004657	0.01819	39.53	303.26
4	negative regulation of leukocyte differentiation (GO:1902106)	0.0004657	0.01819	39.53	303.26
5	antifungal innate immune response (GO:0061760)	0.000003294	0.0003447	22.68	286.24
6	neutrophil activation involved in immune response (GO:0002283)	3.896e-25	1.248e-21	5.05	284.05
7	neutrophil degranulation (GO:0043312)	1.184e-24	1.896e-21	5.01	276.26
8	neutrophil mediated immunity (GO:0002446)	2.876e-24	3.070e-21	4.93	267.14
9	immune response-inhibiting cell surface receptor signaling pathway (GO:0002767)	0.0001119	0.006307	26.39	240.05
10	positive regulation of apoptotic cell clearance (GO:2000427)	0.0001119	0.006307	26.39	240.05

ارتباط موارد گزارش شده به سیستم ایمنی درستی نتایج ما را نشان می دهد.

۳. الف) در بخش transcription برای aml_up_genes داریم:

Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	E2F4 ENCODE	2.184e-42	2.271e-40	6.51	624.53
2	NELFE ENCODE	2.803e-15	7.289e-14	6.05	202.87
3	FOXM1 ENCODE	3.231e-10	5.600e-9	7.86	171.73
4	TAF7 ENCODE	3.065e-18	1.594e-16	4.03	162.59
5	SIN3A ENCODE	6.484e-18	2.248e-16	3.16	124.93
6	STAT5A ENCODE	1.628e-9	2.419e-8	4.45	89.97
7	E2F1 CHEA	2.901e-8	3.352e-7	2.42	41.93
8	E2F6 ENCODE	2.998e-10	5.600e-9	1.87	40.98
9	BCLAF1 ENCODE	1.255e-7	0.000001004	2.34	37.17
10	ZMIZ1 ENCODE	9.860e-8	8.545e-7	2.30	37.13

- E2F4 and AML: [paper](#)

In this study, human clinical data from the Gene Expression Profiling Interactive Analysis (GEPIA) revealed that increased E2F4 expression was associated with poor prognosis in AML patients. Moreover, the experimental results showed that E2F4 was aberrantly overexpressed in human AML patients and cell lines. Depletion of E2F4 inhibited the proliferation, induced the differentiation and suppressed the growth of AML cells in a nude mouse model.

در این مقاله نشان داده شده است که افزایش E2F4 باعث می شود که پیش بینی AML در بیماران کمتر شود. همچنین داده ها و آزمایش های تجربی نشان داده که E2F4 بیش از حد در انسان هایی که مبتلا به AML هستند دیده شده.

- FOXM1 and AML: [paper](#)

FOXM1 upregulation was also observed in blood cancers including ALL and myeloma. Inhibition of FOXM1 reduced proliferation in AML leukemia cell lines. In addition, FOXM1 was reported to contribute to chemoresistance in AML, although the molecular mechanisms have not been determined.

در این مقاله نشان داده شده است که FOXM1 نقش مثبتی در افزایش سلول های سرطانی AML دارد.

و به همین صورت برای aml_down_genes داریم:

Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	SPI1 ENCODE	0.0001592	0.01655	1.58	13.82
2	ESR1 CHEA	0.004494	0.07790	2.45	13.22
3	TP63 CHEA	0.0009436	0.04636	1.52	10.58
4	VDR CHEA	0.02137	0.2779	2.68	10.30
5	SPI1 CHEA	0.001438	0.04636	1.57	10.28
6	USF2 ENCODE	0.002280	0.04742	1.57	9.53
7	RUNX1 CHEA	0.001783	0.04636	1.50	9.51
8	ZKSCAN1 ENCODE	0.03268	0.3558	2.14	7.33
9	IRF8 CHEA	0.03421	0.3558	2.12	7.17
10	USF1 ENCODE	0.01067	0.1585	1.37	6.23

- SPI1 and AML: [paper](#)

A reduction in Spi1 levels or an indirect inhibition of its activity by cooperating factors involved in leukemic transformation causes AML in humans.

در این مقاله نشان داده شده است که کاهش SPI1 یکی از عوامل ابتدا به AML است.

ب) در مقاله‌های زیر نشان داده شده است که inhibition دو فکتور FOXM1 و E2F4 به درمان این بیماری کمک می‌کند.

- FOXM1 inhibition: [paper](#)

the novel autophagy-dependent mode of FOXM1 suppression by STL427944 validates a unique pathway to overcome tumor chemoresistance and improve the efficacy of treatment with conventional cancer drugs.

- E2F4 inhibition: [paper](#)

HLM006474 was identified using a computer-based virtual screen and the known crystal structure of the DNA bound E2F4/DP2 heterodimer. Treatment of multiple cell lines with HLM006474 resulted in the loss of intracellular E2F4 DNA-binding activity as measured by electrophoretic mobility shift assay within hours. Overnight exposure to HLM006474 resulted in down regulation of total E2F4 protein as well as known E2F targets. The effects of HLM006474 treatment on different cell lines varied, but included a reduction in cell proliferation and an increase in apoptosis. HLM006474 induced apoptosis in a manner distinct from cisplatin and doxorubicin.