

Expression data from normal
adults and fetal human thyroids

اختلاف بیان ژن های تیروئید در
کودکی و بزرگسالی

پروژه زیست شناسی سامانه

استاد مربوطه: دکتر مهدی صادقی

دانشجو: زهرا الهیار

9811321006

زیست شناسی سلولی و مولکولی



تیروئید جزء بزرگ‌ترین غدد اندوکرین بوده و از دو لوب تشکیل شده‌است. وزن تیروئید حدود بیست و پنج گرم و اندازه آن در مردان بزرگتر است. اندازه تیروئید در زنان، در دوران قاعدگی و حاملگی افزایش می‌یابد. تیروئید در قسمت مرکزی دارای ایسموس (تنگه) است که سبب اتصال دو لوب تیروئیدی به یکدیگر می‌شود. ممکن است گاهی لوب سومی هم به شکل هرمی از ایسموس یا دو لوب اصلی دیگر بیرون بزند. در بعضی موارد توده‌های کوچکی از جنس بافت تیروئید در اطراف غده مشاهده می‌شود که به غده اصلی اتصالی ندارند و با نام غدد اکسسوری (فرعی) تیروئیدی شناخته می‌شوند.

واحدهای تشکیل دهنده غده تیروئید آسینوس یا فولیکول است و در قسمت مرکزی آن پروتئین‌های کلوئیدی قرار دارد که خود به عنوان انبار ذخیره هورمون تیروئید استفاده می‌شود. به شکل پروانه در جلوی گردن در پایین حنجره و بالای استخوانهای ترقوه، تیروئید دو نوع هورمون ترشح می‌کند به نام های T4 و T3 کار این هورمون ها تنظیم چگونگی مصرف و ذخیره انرژی در بدن است. همچنین غده تیروئید در کنار دو هورمون T3 و T4 هورمونی به نام کلسی‌تونین نیز تولید می‌کند که مسئولیت تنظیم مقدار کلسیم خون بر عهده آن است. در حقیقت تیروئید مسئول تنظیم سوخت و ساز بدن می‌باشد. عملکرد تیروئید توسط غده دیگری در بدن به نام هیپوفیز تنظیم می‌شود.

هیپوفیز در مغز قرار دارد و با ترشح هورمون محرک تیروئید، باعث افزایش ترشح هورمونهای تیروئید می‌شود. تیروئید ممکن است دچار کم کاری یا پرکاری شود. کم کاری تیروئید: کم کاری تیروئید شرایطی است که غده تیروئید قادر به تولید هورمون تیروئید کافی نیست.

پرکاری تیروئید زمانی رخ می‌دهد که غده تیروئید بیش از حد نرمال هورمون ترشح می‌کند. از آنجا که هدف اصلی هورمون تیروئید "تنظیم متابولیسم بدن" است، درواقع مبتلایان به این بیماری علائم مرتبط با کم شدن متابولیسم بدن را دارند

در این پروژه به بررسی میزان تفاوت بیان ژن در افراد بزرگسال طبیعی و جنین و ارزیابی بلوغ و تمایز عملکردی تیروئید در می پردازیم.

نحوه کار به این صورت است که برای مقایسه RNAهای حاصل از تیروئید افراد بالغ و جنین را با روش Affymetrix arrays باهم هیبرید کرده و نتیجه را با بافت های اندام مختلف در انسان بالغ و جنینی مقایسه و بررسی می کنیم .

در این پروژه داده های خام مورد نیاز خود را از پایگاه GEO که، یکی از کاملترین پایگاه های داده برای آزمایشات میکروآری است و همچنین اطلاعاتی چون شرایط آزمایش ، اطلاعات نمونه ها ، اطلاعات شرایط پاتوبیولوژیکی و ... را در بر میگیرد ، به تعداد 10 عدد نمونه با فرمت

Affymetrix Human Genome U133 Plus 2.0 Array [HG-U133_Plus_2] دانلود میکنیم.

داده های مورد نظر را در نرم افزار R وارد میکنیم .

باتوجه به این که هر بافت و سلول عملکرد های متفاوتی دارد در نتیجه این بیان ژن ها میتواند بسیار باهم دیگر متفاوت بوده و مقیاس آماری صحیحی نتوانیم از آن بدست آوریم.

برای اینکه داده ها قابل مقایسه و اصطلاحاً هم رنج شوند قبل از شروع داده های خام را پیش پردازش و نرمال می کنیم.

قبل از نرمال کردن کنترل کیفیت داده ها را بررسی کرده تا از قابل تایید بودن داده ها اطمینان بیابیم و بعد داده ها نرمال و در آخر مجدد کنترل کیفیت بعد از نرمالیزاسیون هم انجام میدهیم جهت اطمینان از نرمال شدن صحیح و قابل قبول داده های مورد نظرمان.

پیش پردازش و کنترل کیفیت داده ها به ترتیب روش های زیر :

1-تصاویر کاذب

2-رسم هیستوگرام

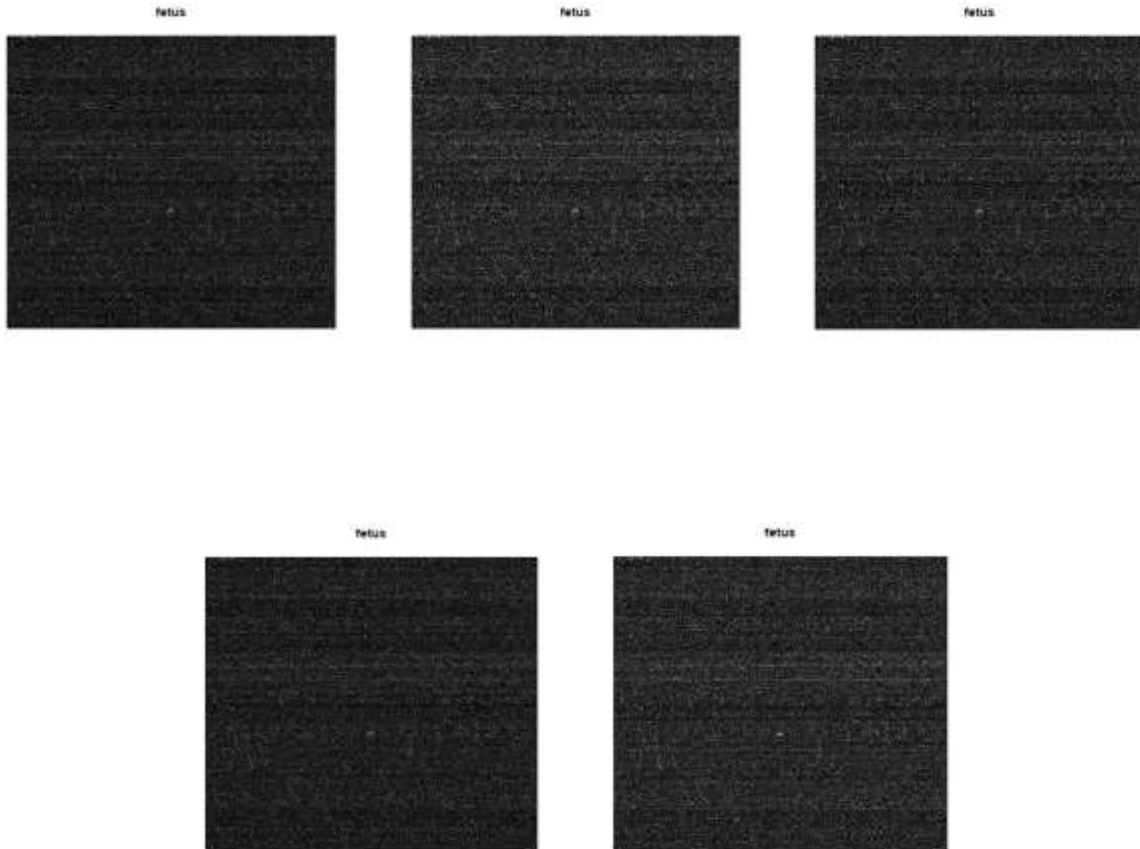
3-رسم boxplot

4-رسم MA plot

به ترتیب هر کدام از این مراحل را برای نمونه های خود یک بار قبل از نرمال کردن و یک بار بعد از نرمال کردن داده ها انجام ایم که تصاویر هرمرحله به ترتیب گزارش شده است.

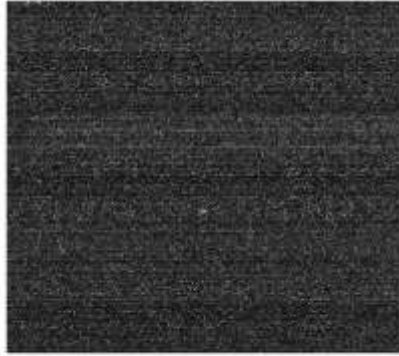
در affymetrix از chip ها برای هیبرید کردن استفاده می کنیم پس برای اطمینان از سالم بودن chip ها میتوانیم تصاویر کاذب داده هام خام تهیه کنیم .

تصاویر کاذب مربوط به بافت های جنینی:



تصاویر کاذب مربوط به بافت های انسان بالغ :

adult



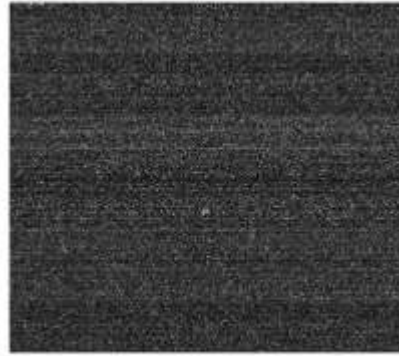
adult



adult



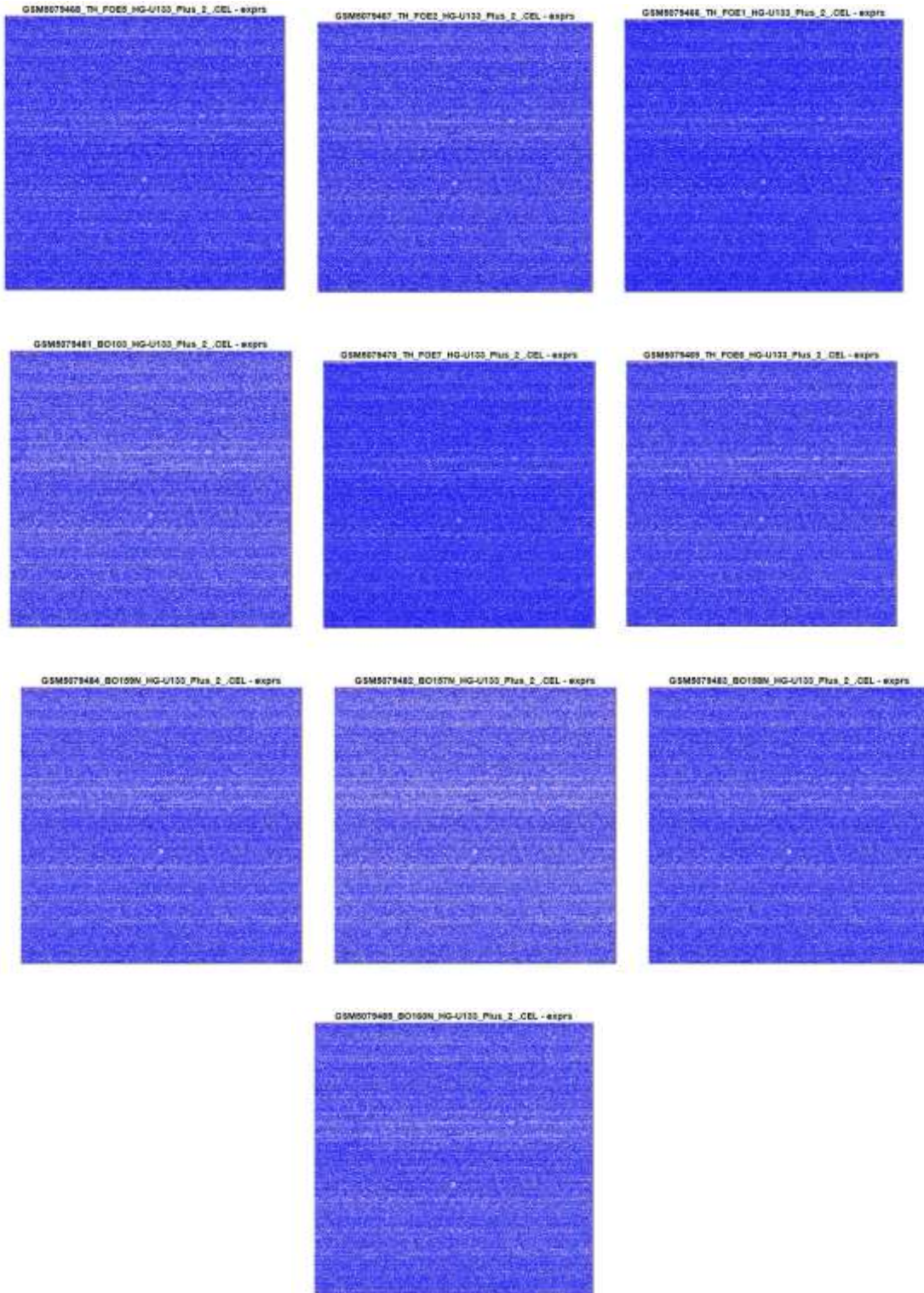
adult



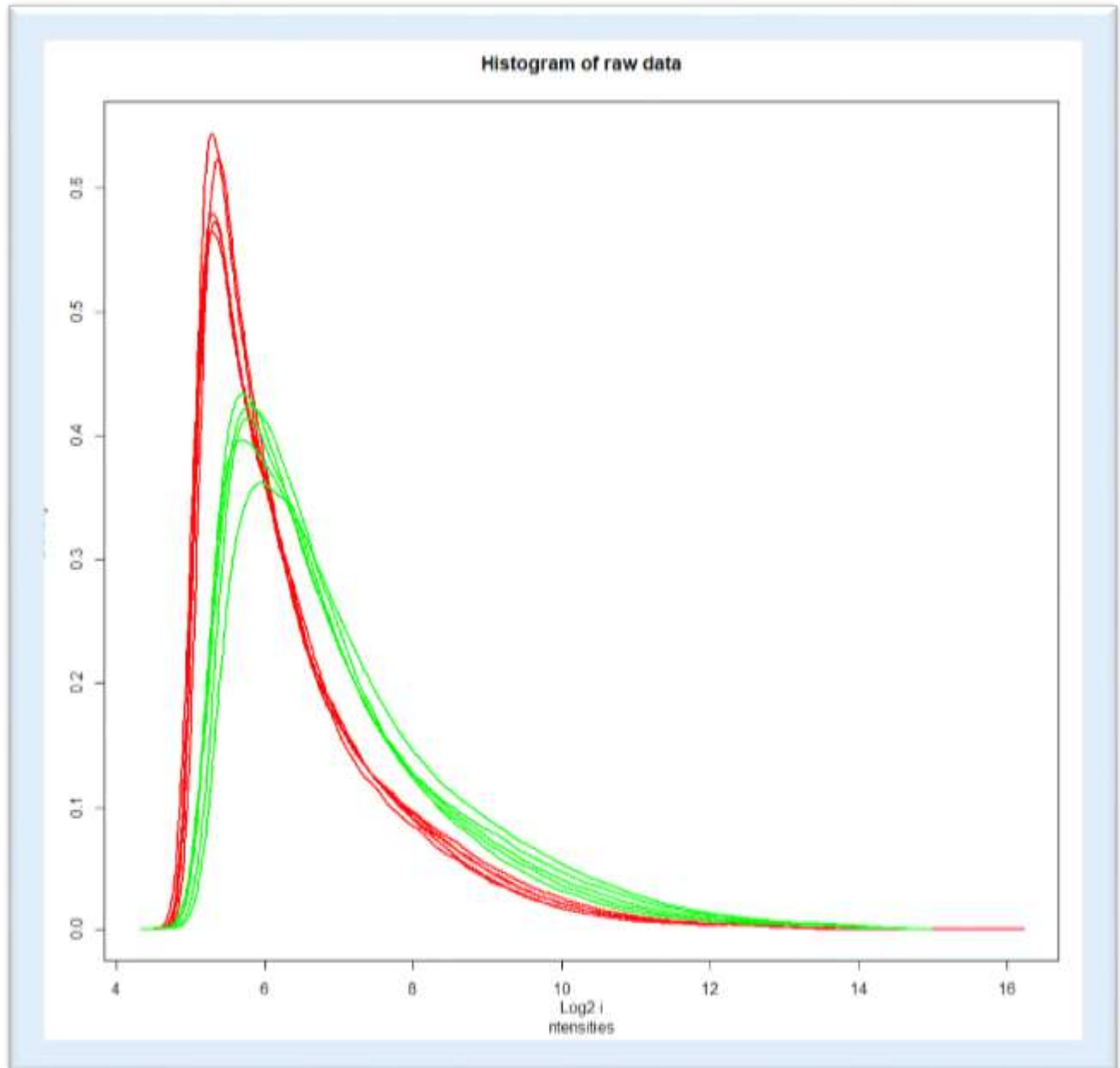
adult



اگر در رسم تصاویر از پکیج oligo استفاده کنیم نتیجه به این صورت خواهد بود:



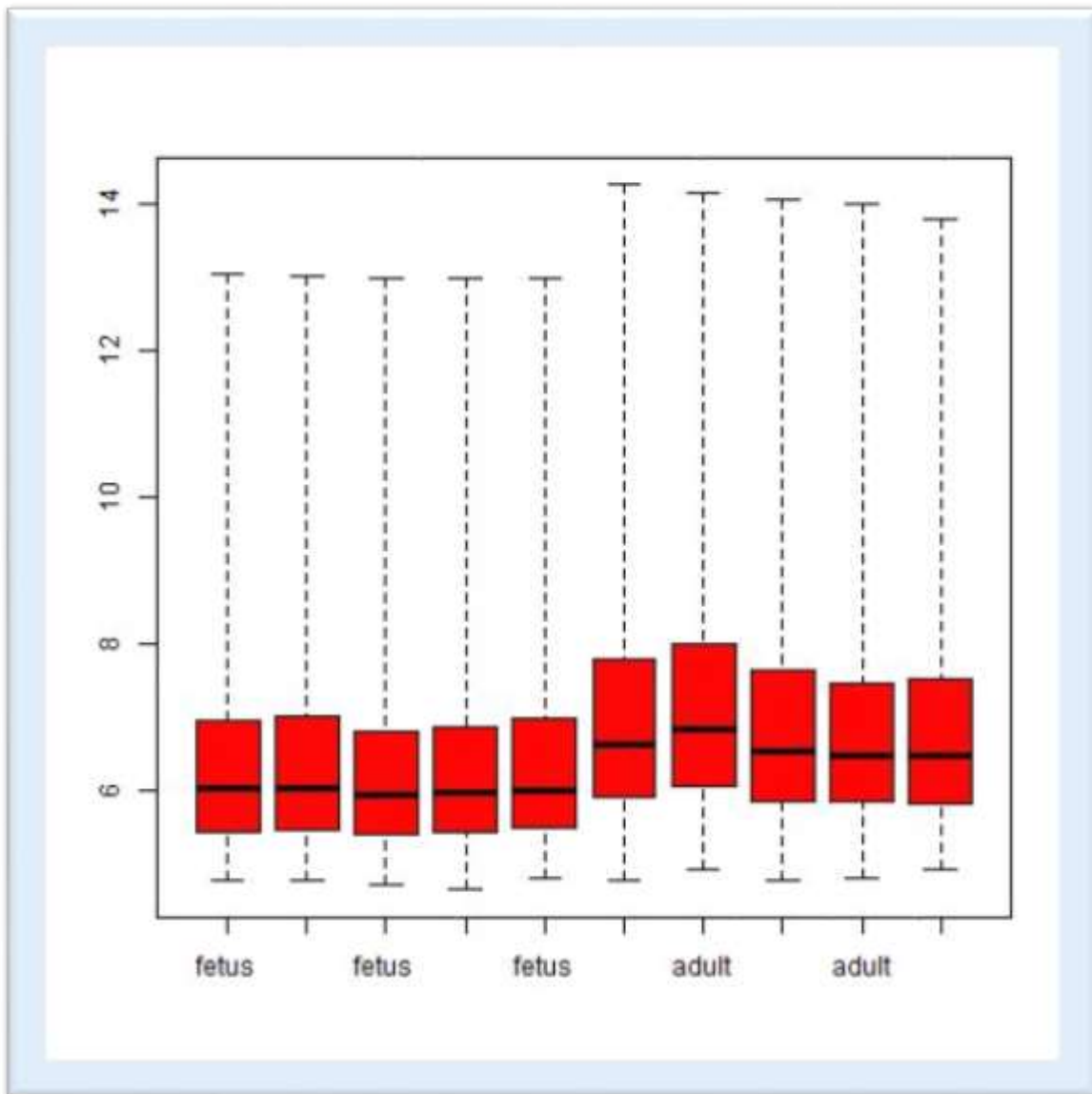
یکی دیگر از روش های کنترل کیفی رسم نمودار هیستوگرام است .
هیستوگرام یک نمودار فراوانی است و میتوان برای هر چیپ آن را رسم کرد.
با ملاحظه نمودار به راحتی متوجه اختلاف بین نمونه ها میشویم.



نمودار بعدی **box plot** است که برای هر نمونه یک باکس رسم می کند و در نمودار محور افقی نمونه ها و محور عمودی مقادیر بیان زن است که در مقیاس لگاریتمی محاسبه شده اند.

در این نمودار مقادیر بیان زن اکثر نمونه ها به حالت یک **box** مشخص شده است و درون **box** نیز مقادیر میانه داده ها مطرح شده است.

درواقع این نمودار به پیدا کردن محل توزیع بیان زن های کمک میکند.



آخرین مرحله کنترل کیفیت قبل از نرمال کردن داده ها رسم نمودار MA plot است

نمودار MA plot از دو مقدار M value و A value تشکیل میشود

این نمودار را برای میکروآرای دو رنگ رسم کرده و به صورت لگاریتم نسبت بیان نمونه ها محاسبه میکنیم.

$$M = \log_2 \frac{fetus}{adult}$$

تفسیر جواب حاصل به 3 صورت خواهد بود:

اگر مقدار لگاریتم ما برابر 1 باشد: افزایش بیان

اگر مقدار لگاریتم ما برابر -1 باشد: کاهش بیان

اگر مقدار لگاریتم ما برابر 0 باشد : عدم وجود تفاوت بیان

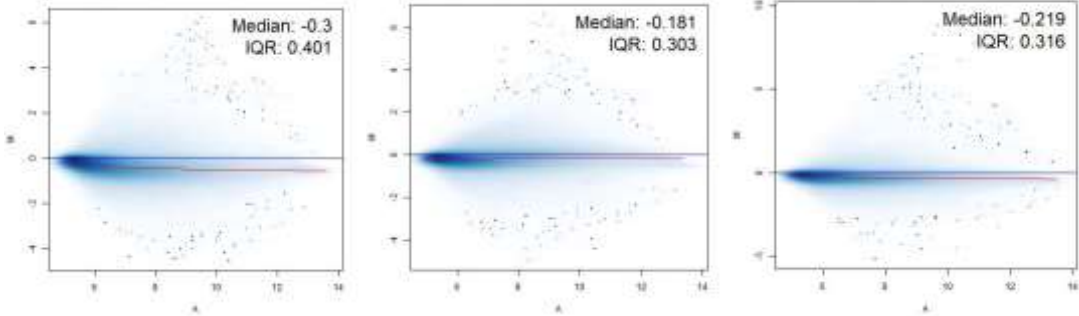
پارامتر M value بیانگر تفاوت بیان ژن در نمونه هاست و پارامتر A value بیانگر مقدار کلی بیان ژن در نمونه هاست .

برای محاسبه A value از رابطه زیر استفاده می شود:

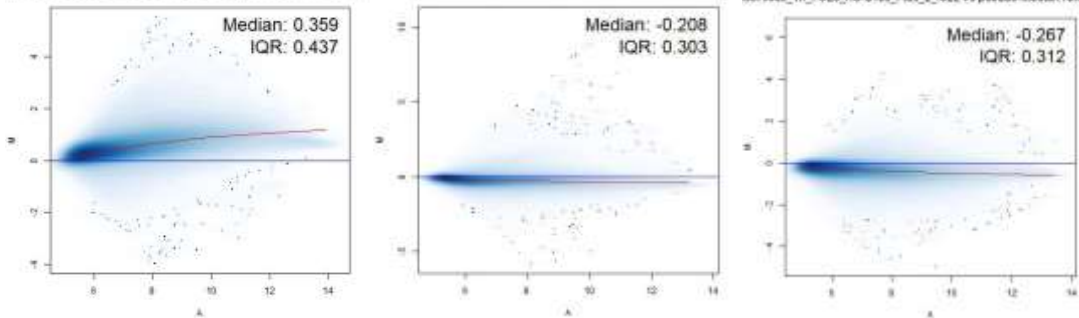
$$A = \frac{1}{2} (\log_2 fetus + \log_2 adult)$$

رسم MA plot با پکیج الیگو

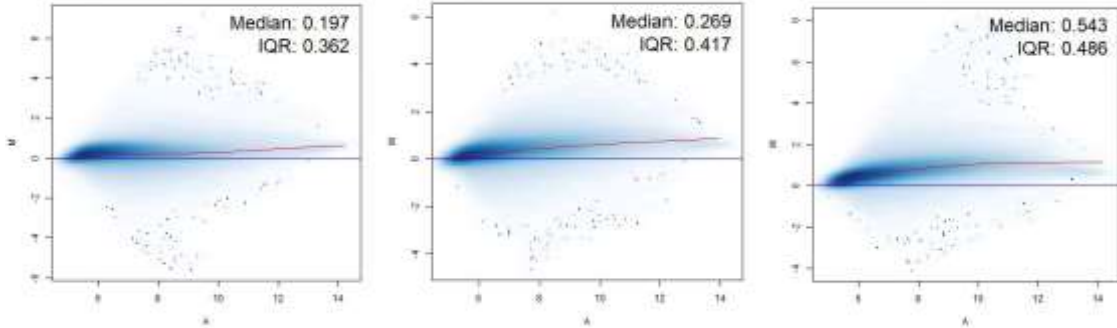
M079468_TH_F0E8_HG-U133_Plus_2_CEL vs pseudo-median referen M079467_TH_F0E2_HG-U133_Plus_2_CEL vs pseudo-median referen M079466_TH_F0E1_HG-U133_Plus_2_CEL vs pseudo-median referen



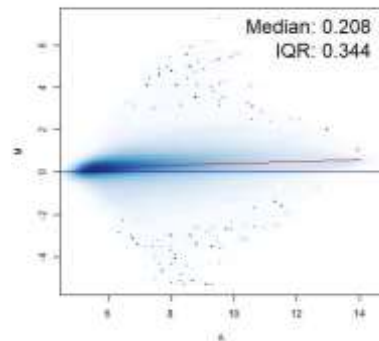
M079461_BO103_HG-U133_Plus_2_CEL vs pseudo-median referen M079470_TH_F0E7_HG-U133_Plus_2_CEL vs pseudo-median referen M079469_TH_F0E6_HG-U133_Plus_2_CEL vs pseudo-median referen



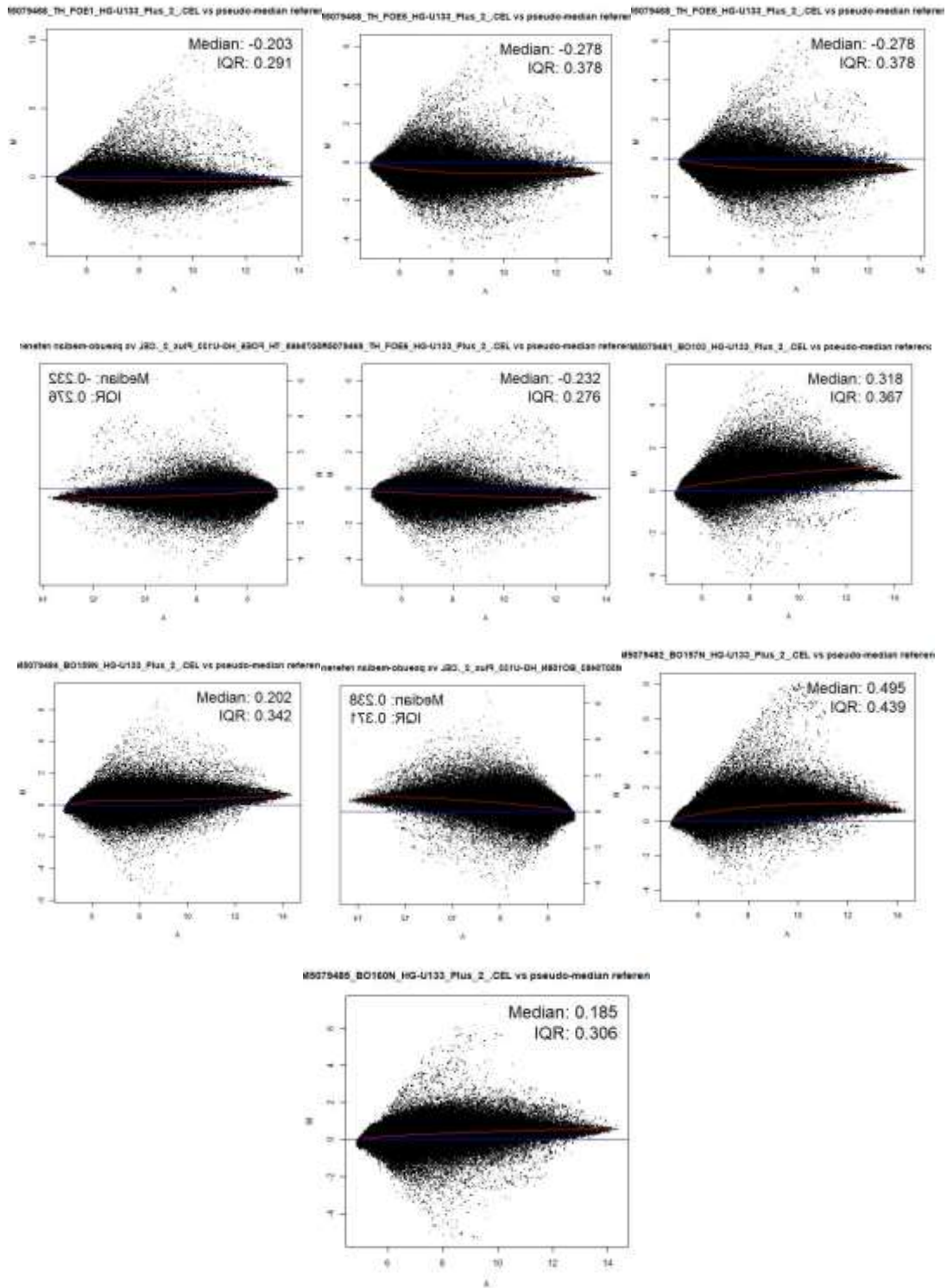
M079464_BO188N_HG-U133_Plus_2_CEL vs pseudo-median referen M079463_BO188N_HG-U133_Plus_2_CEL vs pseudo-median referen M079462_BO187N_HG-U133_Plus_2_CEL vs pseudo-median referen



M079465_BO160N_HG-U133_Plus_2_CEL vs pseudo-median referen



رسم MA plot با پکیج affy



بعد از انجام کنترل کیفیت و اطمینان از کیفیت خوب داده ها، داده های خام را با طی کردن مراحل زیر نرمال کردیم تا داده ها قابل مقایسه شده و خطاهای تکنیکی را حذف کرده باشیم.

1-Background correction

2-Quantile normalization

3-summarization

نرمال کردن داده ها باعث توزیع نرمال و راحت تر شدن مقایسه آن ها باهم می شود

مرحله اول باعث حذف نویز ها و بهتر نشان داده شدن داد های ما میشود به این صورت که در هر chip اگر ژنی وجود داشته باشد و با مکملش هیبرید شده باشد نور فلورسنت از خودمنتشر میکند و میزان بیان ژن با میزان نور رابطه مستقیم دارد.

در مرحله بعد با ماتریس رتبه دهی مقدار عددی این بیان ژن ها را هم در یک رنج قرار میدهیم به طور که اگر نمودارش را رسم کنیم توزیع به صورت زنگوله ای خواهد شد و در آخرین مرحله اطلاعات هر 16 الی 22 pm را خلاصه و یک عدد واحد بدست می آوریم.

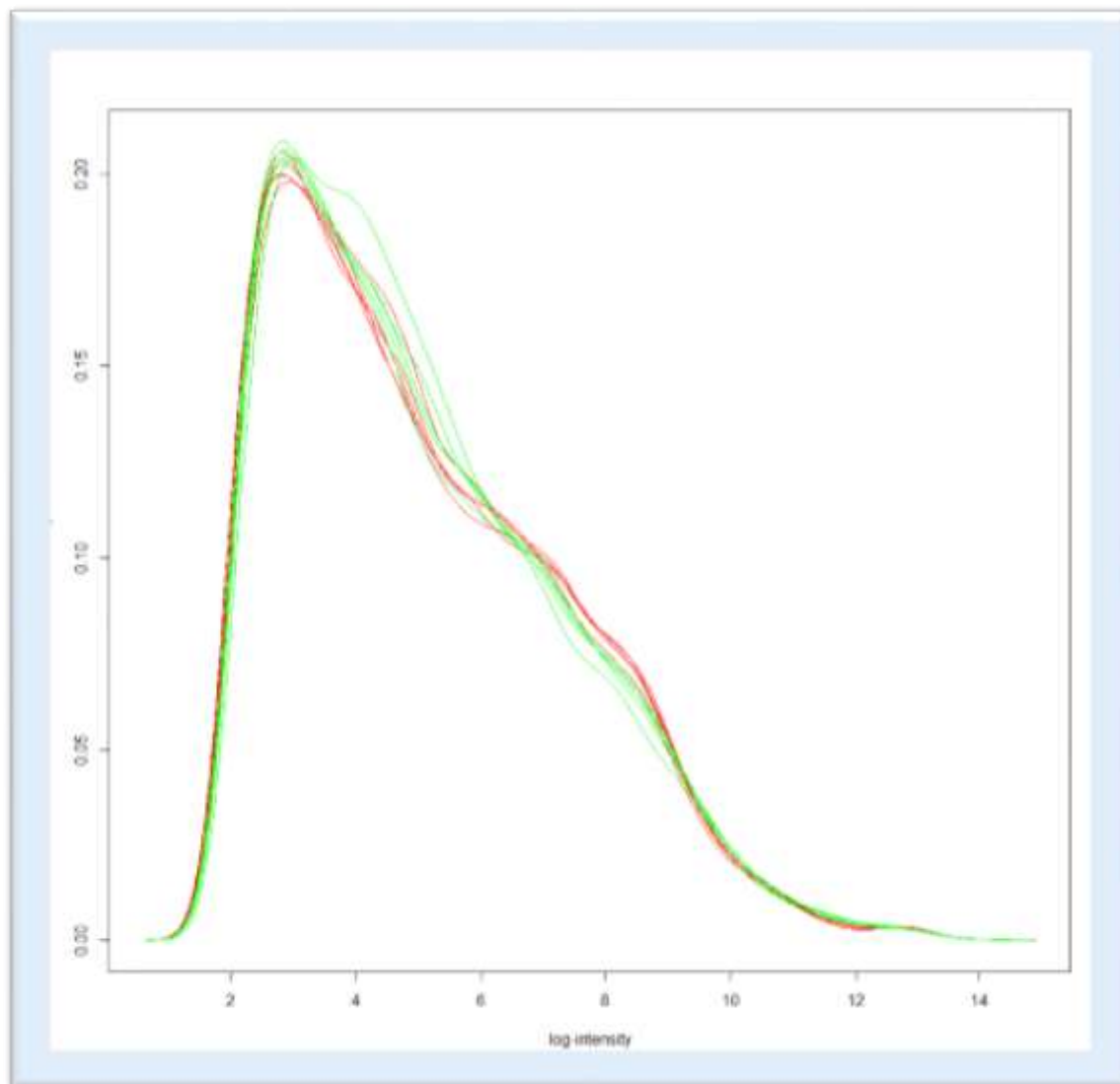
بعد از نرمال کردن داده مجدد کنترل کیفیت انجام داده و مجدد سه نمودار زیر را برای داده هایمان رسم کردیم :

هیستوگرام

Box plot

MA plot

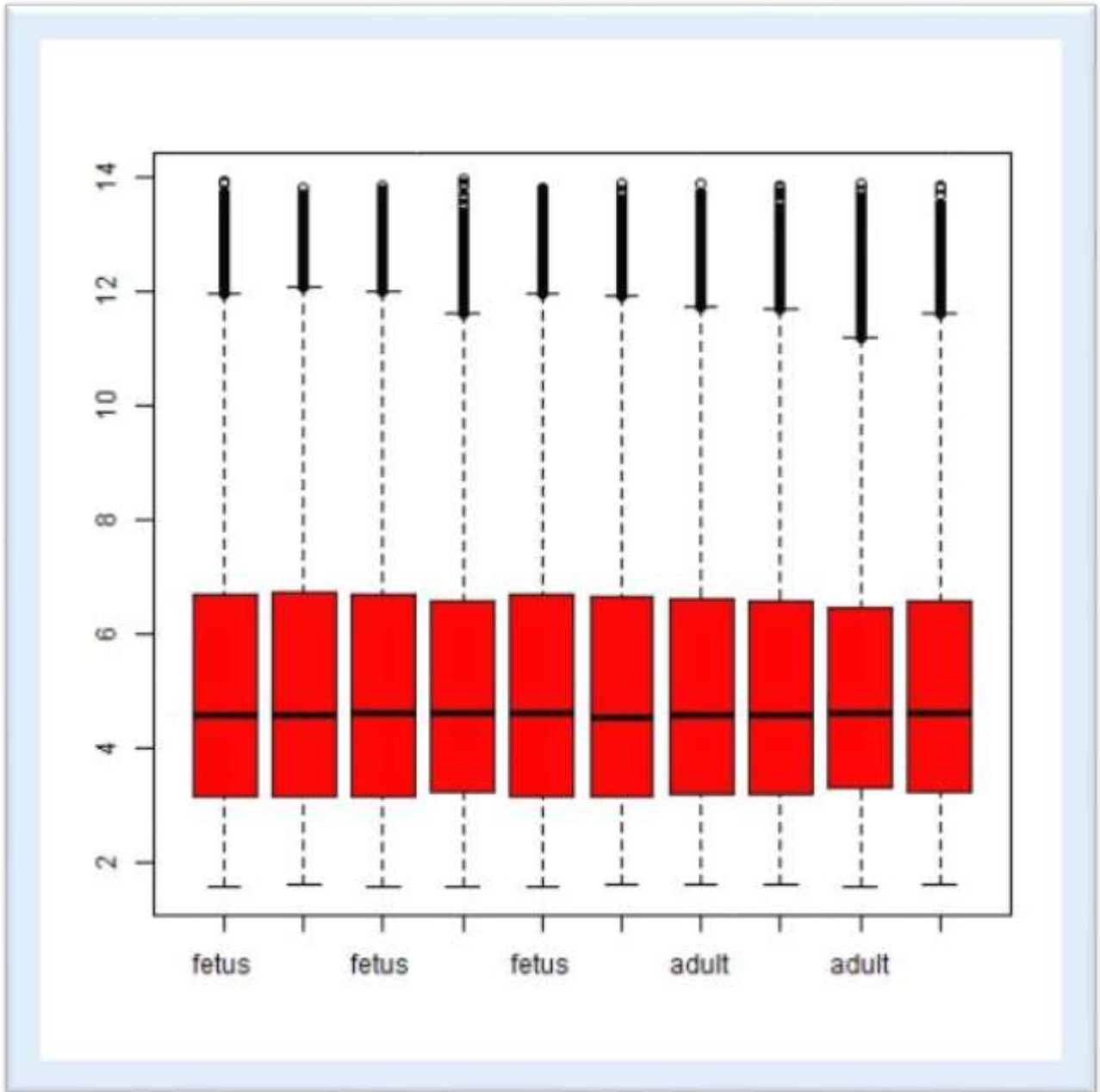
نمودار هیستوگرام برای داده های نرمال شده :



همینطور که در تصویر هم قابل ملاحظه هست ، داده ها تقریباً در یک رنج قرار گرفته اند و نمودار برای هر کدام از داده ها بر روی هم منطبق است و مقایسه آنها خیلی راحت تر از داده های نرمال نشده است.

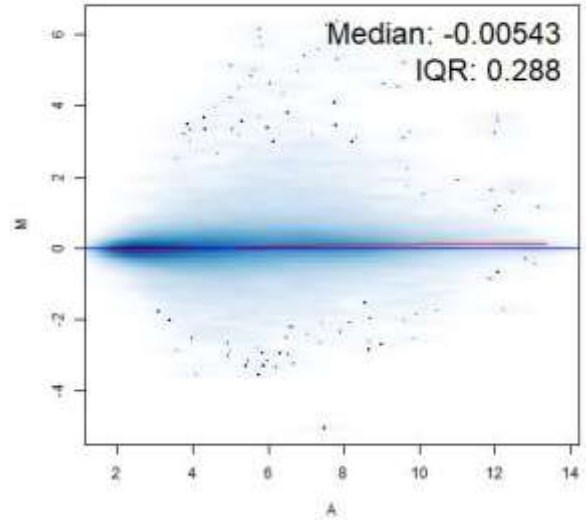
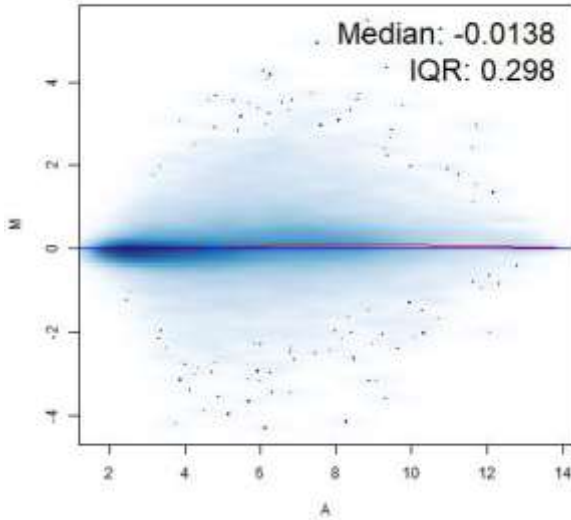
هرچه نمونه های ما شرایط شبیه تری داشته باشند این اختلاف بین آنها کمتر خواهد بود و بیشترین میزان اختلاف بیان ژن در بدترین حالت برابر 5٪ است.

نمودار boxplot برای داده های نرمال شده :

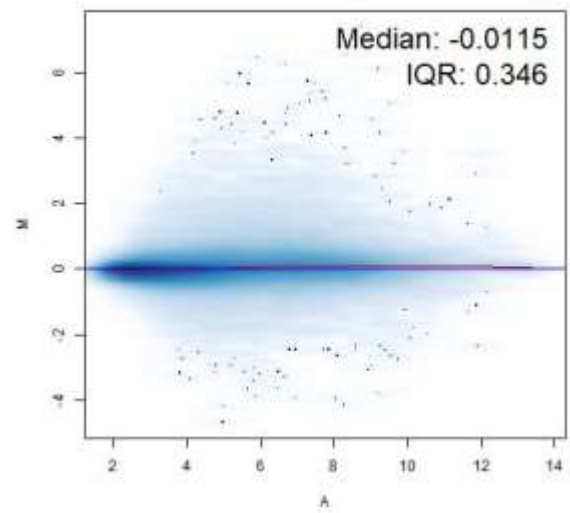
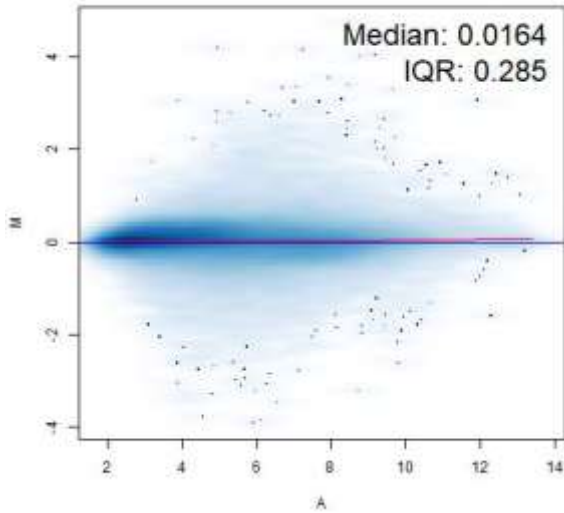


نمودار MA plot برای داده های نرمال شده با پکیج الیگو:

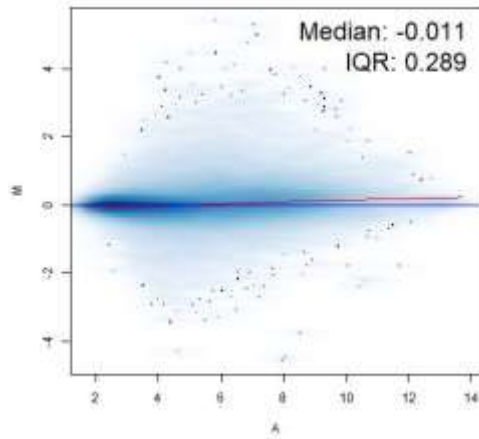
15079467_TH_FOE2_HG-U133_Plus_2_CEL vs pseudo-median referer15079466_TH_FOE1_HG-U133_Plus_2_CEL vs pseudo-median referer



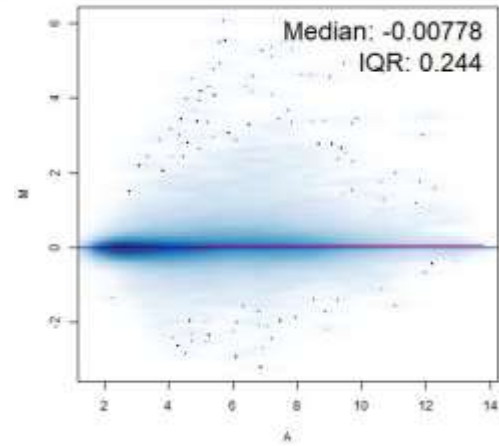
15079469_TH_FOE6_HG-U133_Plus_2_CEL vs pseudo-median referer15079468_TH_FOE5_HG-U133_Plus_2_CEL vs pseudo-median referer



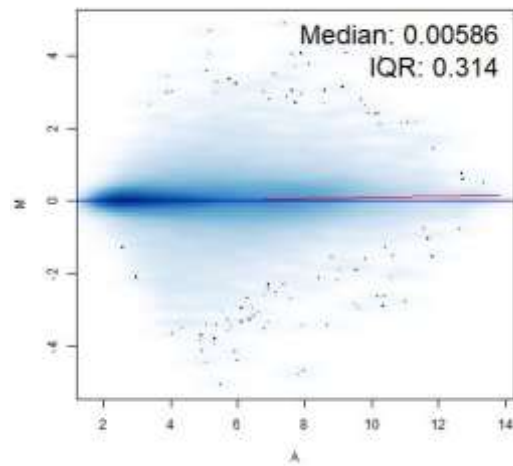
M5079481_BO103_HG-U133_Plus_2_CEL vs pseudo-median referenc



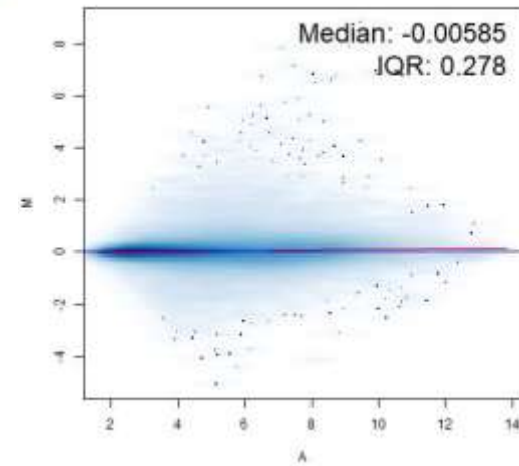
M5079470_TH_FOE7_HG-U133_Plus_2_CEL vs pseudo-median referer



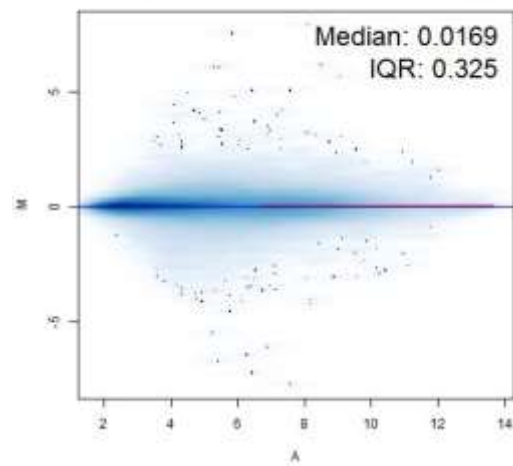
M5079483_BO158N_HG-U133_Plus_2_CEL vs pseudo-median referen



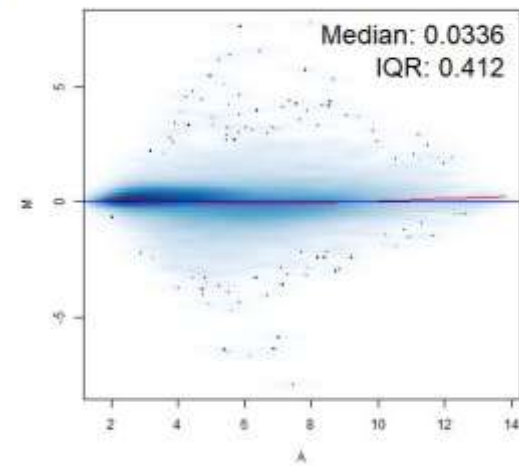
M5079482_BO157N_HG-U133_Plus_2_CEL vs pseudo-median referen



M5079485_BO160N_HG-U133_Plus_2_CEL vs pseudo-median referen

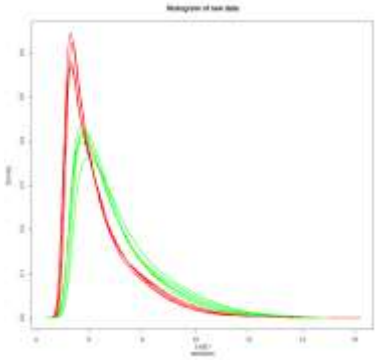
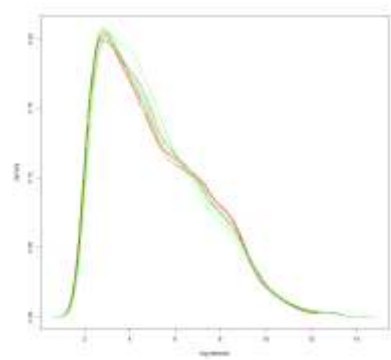
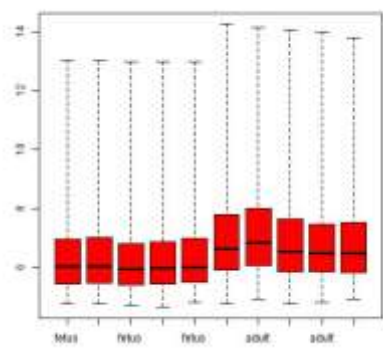
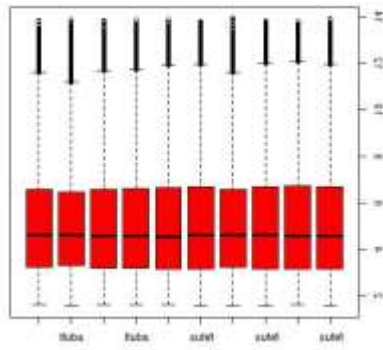
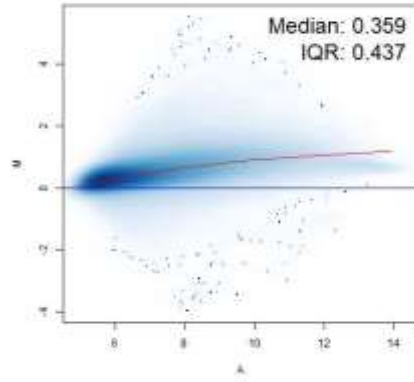
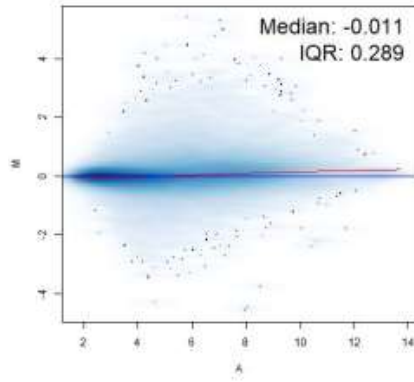


M5079484_BO159N_HG-U133_Plus_2_CEL vs pseudo-median referen



شمای کلی از داده و مقایسه بین حالت نرمال شده و نرمال نشده :

M5079481_BO103_HG-U133_Plus_2_CEL vs pseudo-median referencM5079481_BO103_HG-U133_Plus_2_CEL vs pseudo-median referenc



بعد از نرمال کردن داده های مورد نظرمان داده ها را بررسی و ژن هایی که طبق تحلیل مان دارای تفاوت در بیان ژن بودند را پیدا کردیم و این ژن ها به شامل:

تعداد ژن های دارای بیان بدست آمده 1429 عدد می باشد که در فایل sheet1 ارائه خواهد شد

MCM4	HS6ST2	EGR2	AGPAT5	SERPINH1	HLA-DQA1
LRP1B	SULT1E1	CPXM1	F2RL2	HS6ST2	PTPRC
ST6GALNAC2	SEPTIN6	PSMB9	MN1	COL4A6	CPEB3
DDIAS	CDCA2	HUNK	MEGF10	TNNC2	TSHZ2
ROR1	RGS16	MX1	OSR2	RASGRF2	NLGN4X
TTN	CTSS	SEPTIN6	ESR1	STEAP4	FBXL16
CAPS2	CD69	DEPDC7	NR4A2	CARD16	TUBB2B
CLU	IFI44L	MARCKS	NEBL	ID2-AS1	PFKFB2
CCNL1	IGFBP7	LRRK2	PARPBP	MAD2L1	SPDL1
MEOX1	ETV1	CLIC5	PTPRC	SPOCK3	GPR183
HEMGN	CENPF	CLU	ADAM12	PTPRB	OGN
HCP5	ASPM	CCL5	RNASEH2A	BHMT2	CCNB2
FAM171A1	TUSC3	GGH	CNMD	STMN2	COL3A1
BIRC3	ADGRL3	CEBPD	SLC10A4	WIF1	SHCBP1
PDK4	CD48	RGS1	HPGD	ATP13A4	CIP2A
TUSC3	HMGB3	SERPINB1	KRT7	SLC5A3	CYS1
ZNF595	TUBB2A	IFI6	SP2-AS1	NMU	TNNI2
STEAP4	PRKX	WASF1	MCM10	COL21A1	CDC6
PTPRD	NREP	IGKC	PAFAH1B3	NCAPG	LINC-PINT
HLA-DQB1	WDR72	TXNL1	GPR37	TNFRSF11B	MCM6
SLC4A11		CCN2+B3B362:B373A	CDH11	CDH11	JPT1
CDH2	NUF2	LRRC17	PEG10	PEG10	CDC45
CRYBG1	NEBL	RBP7	MUC15	MUC15	P2RY14
HSD11B1	AGR2	CENPU	PDGFRA	PDGFRA	SOX8
MT1G	CDCA8	NAV3	PIEZO2	PIEZO2	SEPTIN6
TIAM1	GTSE1	CLEC2B	PXDN	PXDN	FBXO5
GNAS	TMEM139	GYPA	GBP2	GBP2	RARRES1
AGTR1	TOX	XAF1	PLP1	PLP1	GBP3
DEFA1	CCR2	MCM10	ENC1	ENC1	THBS1
NID2	DLK1	TRIM59	DGKI	DGKI	LRRC3B
FBN3	GPSM1	ATP1B4	CBX2	CBX2	ARHGAP28
SNORD3B-1	DPT	CDH2	HMGA2	HMGA2	FOXG1
CFH	CDKL2	FAM167A	CDH11	CDH11	NUSAP1
GLRB	MYH8	GREB1	GREM1	GREM1	H2AX
F8	EGFL6	CDCA3	IL15	IL15	RAD51
CGA	AHSP	OSR1	CLCN5	CLCN5	ELOVL6
			ITGA2	ITGA2	COL1A2

RHOBTB3	CDC6	ARL4D	TMSB15A	POSTN	HAS2
CAV2	MT1E	TENM2	MT1X	ADGRL3	POSTN
SLPI	HLF	STMN1	RAD51AP1	CHEK1	SLC26A4
MYH3	ABCA8	IGH	BRCA2	NACC2	TMEM97
RNF175	NEAT1	ATF3	KPNA2	PHF19	NUDT11
POGLUT2	COL3A1	MYL1	DPYSL3	CA10	PCDH8
LRRC2	TMEM171	EBF3	CDH3	ANXA9	EBF3
IFITM2	GPRC5A	FGL2	DBF4	ADH1B	FGL2
TSPAN8	DNASE1L3	WNT5A	BARD1	SNAP25	WNT5A
PAK3	NT5E	ZGRF1	MECOM	SUCNR1	ZGRF1
ROR1	PTN	TMEM267	0	ARL4D	TMEM267
NALCN	HLA-DPA1	DNAJC12	BCAT1	TENM2	DNAJC12
NAV2	COL4A2	PLA2R1	TMEM132C	STMN1	PLA2R1
EDNRA	NRK	PLA2R1	ADAMTS2	IGH	PLA2R1
UHRF1	MMP16	HLA-F	HHATL	ATF3	HLA-F
GPSM2	HOXB3	MAN1C1	COL12A1	MYL1	MAN1C1
CDC7	DHFR	LRRN1	OIP5	OSR1	LRRN1
COL4A3	ZBTB20	COL4A4	C1QTNF4	NT5E	COL4A4
PLN	HOXA3	CRYBG3	HSPB3	UBTD2	CRYBG3
SGO2	GINS3	IGLL5	FCGR2B	LACC1	IGLL5
CCNE2	JPT1	GADD45B	TARP	MYEF2	GADD45B
PHF19	PLN	RTP4	CHEK1	FAM72C	RTP4
BARD1	TRIB1	ADAMTS18	CD59	HBA1	ADAMTS18
BHMT2	RHOBTB3	MYL11	HLF	LOC102724701	MYL11
MYEF2	EZH2	DIAPH3	BMP8A	ALAS2	DIAPH3
DEPDC1B	SFRP1	TCF19	ZMAT4	CDH11	TCF19
INMT	ANXA9	RHOBTB1	IGIP	FOSL2	RHOBTB1
HELLS	ASPA	LOC100288637	NRXN3	APOLD1	LOC100288637
GBP1	NACC2	CYP27C1	SCNN1A	ORC6	CYP27C1
NTM	PHF19	DACH1	CXCL14	COBL	DACH1
CLIC6	CA10	CD59	IL15	HS6ST2	CD59
C3orf52	ANXA9	COL5A1	UBE2T	SMC2	COL5A1
ANO5	ADH1B	LIX1	SLC26A4-AS1	CREBRF	LIX1
NSD2	SNAP25	IFITM1	BCAT1	GAS2	IFITM1
ZNF883	SUCNR1	NEAT1	LPAR4	ZBTB16	
		TRIP13	IGF2		

functional annotation

affymetrix ID را که از پکیج lima بدست آوریدم را با استفاده از سایت Divid به اسم ژن تبدیل میکنیم تا مسیرهایی که ژن مورد نظرمان در آن درگیر است را پیدا کنیم. در این سایت از بخش های MF fat ,BP fat , CC fat , اطلاعات و لیست ژن های درگیر در هر بخش را بدست آوردیم.

*برای هریک از تعداد ژن ها بسیار زیاد بود و در اینجا فقط تعدادی ژن به عنوان نمونه قرارداد شده است و فایل کامل پیوست می گردد. shee2-sheet10

Biological prosses

از سایت DAVID مسیر های مربوط به ژن مورد نظر را جستجو کرده و به این نتیجه رسیدیم که این ژن ها در مسیرهای مانند چرخه سلولی، مرگ سلولی ، پاسخ به محرکهای شیمیایی، ارتباطات سلولی، جداسازی کروموزم ها ،متابولیسم پروتئین ها وسیگنالینگ دخیل هستند. برای هر کدام به عنوان نمونه چند ژن قرارداد شده است.

لیست برخی از ژن های درگیر در چرخه سلولی بر پایه BP-FAT

Gene Name
15-hydroxyprostaglandin dehydrogenase(HPGD)
ATP binding cassette subfamily B member 1(ABCB1)
BLM RecQ like helicase(BLM)
BMP/retinoic acid inducible neural specific 1(BRINP1)
BMP/retinoic acid inducible neural specific 3(BRINP3)

ژن های درگیر در مرگ سلولی

Gene Name
15-hydroxyprostaglandin dehydrogenase(HPGD)
5-hydroxytryptamine receptor 2B(HTR2B)
BMP/retinoic acid inducible neural specific 1(BRINP1)
BRCA1 associated RING domain 1(BARD1)
BRCA2 DNA repair associated(BRCA2)

cellular component assembly

Gene Name
BCL11 transcription factor A(BCL11A)
BLM RecQ like helicase(BLM)
BRCA1 interacting helicase 1(BRIP1)
C-X-C motif chemokine ligand 12(CXCL12)
CD36 molecule(CD36)

cellular response to chemical stimulus

Gene Name
15-hydroxyprostaglandin dehydrogenase(HPGD)
ADAM metallopeptidase with thrombospondin type 1 motif 3(ADAMTS3)
ATP binding cassette subfamily B member 1(ABCB1)
BCL11 transcription factor A(BCL11A)
BLM RecQ like helicase(BLM)

regulation of cell communication

Gene Name
5-hydroxytryptamine receptor 2B(HTR2B)
6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 2(PFKFB2)
ADAM metallopeptidase with thrombospondin type 1 motif 3(ADAMTS3)
BMP and activin membrane bound inhibitor(BAMBI)
C-C motif chemokine ligand 4(CCL4)

regulation of cellular component organization

Gene Name
ATPase family AAA domain containing 2(ATAD2)
BCL11 transcription factor A(BCL11A)
BMP and activin membrane bound inhibitor(BAMBI)
BORA aurora kinase A activator(BORA)
BUB1 mitotic checkpoint serine/threonine kinase B(BUB1B)

regulation of chromosome segregation

Gene Name
15-hydroxyprostaglandin dehydrogenase(HPGD)
5-hydroxytryptamine receptor 2B(HTR2B)
ATP binding cassette subfamily B member 1(ABCB1)
BMP and activin membrane bound inhibitor(BAMBI)
BRCA1 interacting helicase 1(BRIP1)

regulation of protein metabolic process

Gene Name
5-hydroxytryptamine receptor 2B(HTR2B)
BLM RecQ like helicase(BLM)
BORA aurora kinase A activator(BORA)
BRCA1 associated RING domain 1(BARD1)
BUB1 mitotic checkpoint serine/threonine kinase B(BUB1B)

regulation of signaling

Gene Name
5-hydroxytryptamine receptor 2B(HTR2B)
6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 2(PFKFB2)
ADAM metallopeptidase with thrombospondin type 1 motif 3(ADAMTS3)
BMP and activin membrane bound inhibitor(BAMBI)
C-C motif chemokine ligand 4(CCL4)

Cellular compement

این بخش نشان دهنده محل و موقعیت مکانی در سلول است که برای ژن مورد نظر ما بخش هایی مانند کروموزوم، سیتوزول، اندوپلاسم و نوکلئوپلاسم مشاهده شده است. sheet 11-sheet15

Chromosome

Gene Name
BCL11 transcription factor A(BCL11A)
BLM RecQ like helicase(BLM)
BRCA2 DNA repair associated(BRCA2)
BTB domain and CNC homolog 2(BACH2)
BUB1 mitotic checkpoint serine/threonine kinase B(BUB1B)

Cytosol

Gene Name
15-hydroxyprostaglandin dehydrogenase(HPGD)
5'-nucleotidase ecto(NT5E)
6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 2(PFKFB2)
BCL11 transcription factor A(BCL11A)
BLM RecQ like helicase(BLM)

endoplasmic reticulum

Gene Name
1-acylglycerol-3-phosphate O-acyltransferase 5(AGPAT5)
ADAM metallopeptidase with thrombospondin type 1 motif 9(ADAMTS9)
ATP binding cassette subfamily A member 8(ABCA8)
ATPase 13A4(ATP13A4)
BMP/retinoic acid inducible neural specific 1(BRINP1)

extracellular region part

Gene Name
15-hydroxyprostaglandin dehydrogenase(HPGD)
5'-nucleotidase ecto(NT5E)
ADAM metallopeptidase with thrombospondin type 1 motif 18(ADAMTS18)
ADAM metallopeptidase with thrombospondin type 1 motif 2(ADAMTS2)
ADAM metallopeptidase with thrombospondin type 1 motif 3(ADAMTS3)

Nucleoplasm

Gene Name
15-hydroxyprostaglandin dehydrogenase(HPGD)
5'-nucleotidase ecto(NT5E)
5-hydroxytryptamine receptor 2B(HTR2B)
6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 2(PFKFB2)
ADAM metallopeptidase domain 12(ADAM12)

Molecular function

این بخش عملکرد سلول و مسیرهایی که به عملکرد مربوط هستند را بررسی میکند. ژن های که در عملکرد سلول نقش داشتند به ترتیب زیر می باشد: sheet15-sheet21

anion binding

Gene Name
15-hydroxyprostaglandin dehydrogenase(HPGD)
5'-aminolevulinate synthase 2(ALAS2)
6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 2(PFKFB2)
ADAM metallopeptidase with thrombospondin type 1 motif 3(ADAMTS3)
ADP ribosylation factor like GTPase 4D(ARL4D)

Cation binding

Gene Name
5'-nucleotidase ecto(NT5E)
5-hydroxytryptamine receptor 2B(HTR2B)
ADAM metallopeptidase domain 12(ADAM12)
ADAM metallopeptidase with thrombospondin type 1 motif 18(ADAMTS18)
ADAM metallopeptidase with thrombospondin type 1 motif 2(ADAMTS2)

Identical protein binding

Gene Name
15-hydroxyprostaglandin dehydrogenase(HPGD)
5'-nucleotidase ecto(NT5E)
BCL11 transcription factor A(BCL11A)
BLM RecQ like helicase(BLM)
BRCA1 associated RING domain 1(BARD1)

Ion binding

Gene Name
15-hydroxyprostaglandin dehydrogenase(HPGD)
5'-aminolevulinate synthase 2(ALAS2)
5'-nucleotidase ecto(NT5E)
5-hydroxytryptamine receptor 2B(HTR2B)
6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 2(PFKFB2)

metal ion binding

Gene Name
5'-nucleotidase ecto(NT5E)
ADAM metallopeptidase domain 12(ADAM12)
ADAM metallopeptidase with thrombospondin type 1 motif 18(ADAMTS18)
ADAM metallopeptidase with thrombospondin type 1 motif 2(ADAMTS2)
ADAM metallopeptidase with thrombospondin type 1 motif 3(ADAMTS3)

small molecule binding

Gene Name
15-hydroxyprostaglandin dehydrogenase(HPGD)
5'-aminolevulinate synthase 2(ALAS2)
5'-nucleotidase ecto(NT5E)
6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 2(PFKFB2)
ADP ribosylation factor like GTPase 4D(ARL4D)

kEG pathway

sheet22-sheet26

از این پایگاه داده اسفاده می کنیم تا ژن های درگیر در یک مسیر مورد نظر را بیابیم.

در واقع از لیست ژن ها به مسیرهای پیام رسانی میرسیم .

به این پایگاه داده در اصطلاح مبدا شناسی هم گفته میشود.

مسیر های زیر در پایگاه KEG بدست آمده است که خروجی هر کدام به صورت جداول جداگانه در اکسل پیوست می گردد:

- 29 ژن در چرخه سلولی

- 18 ژن در سن سلول

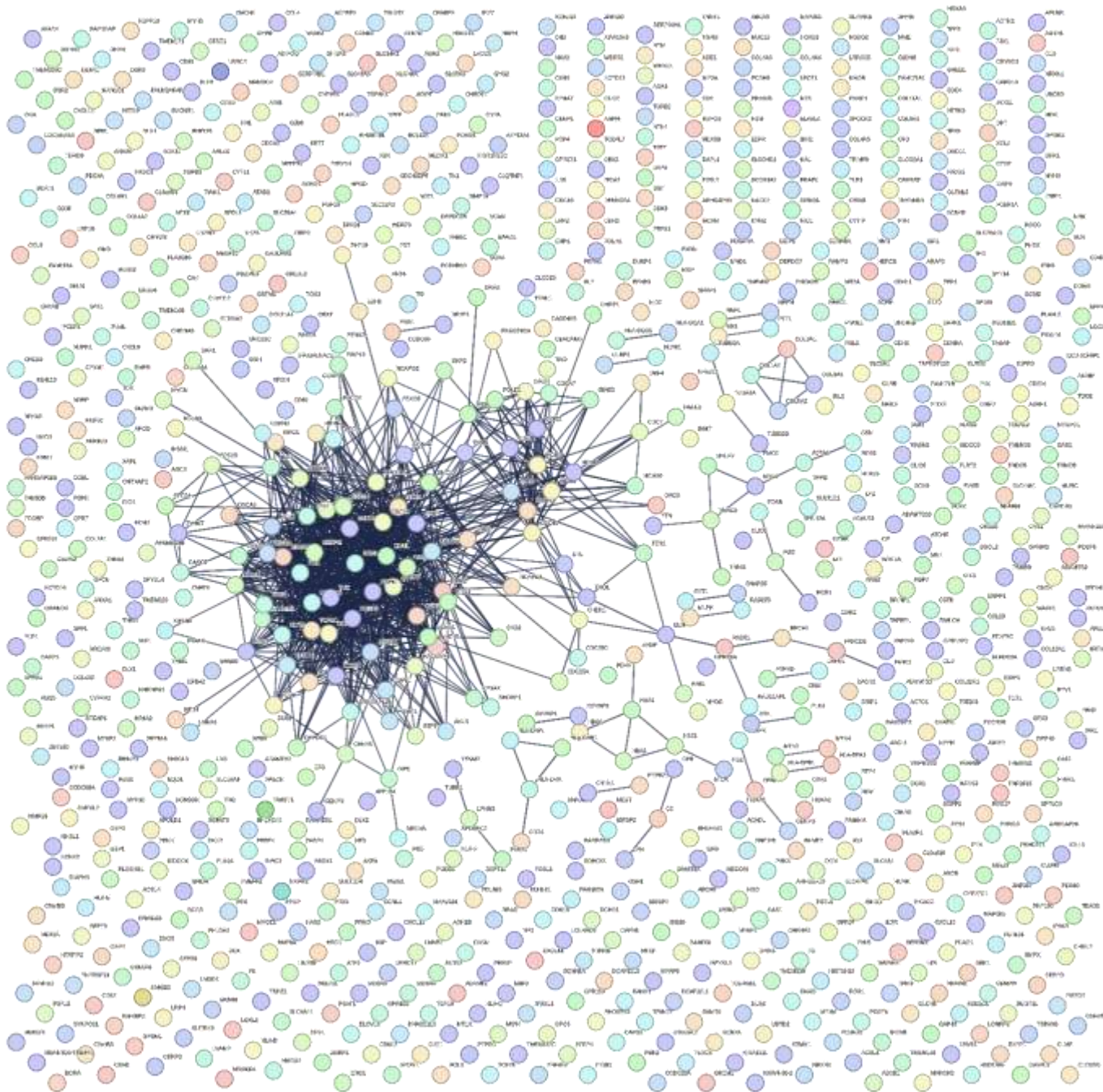
- 23 ژن در چسبندگی سلول

- 24 ژن در فاگوزوم

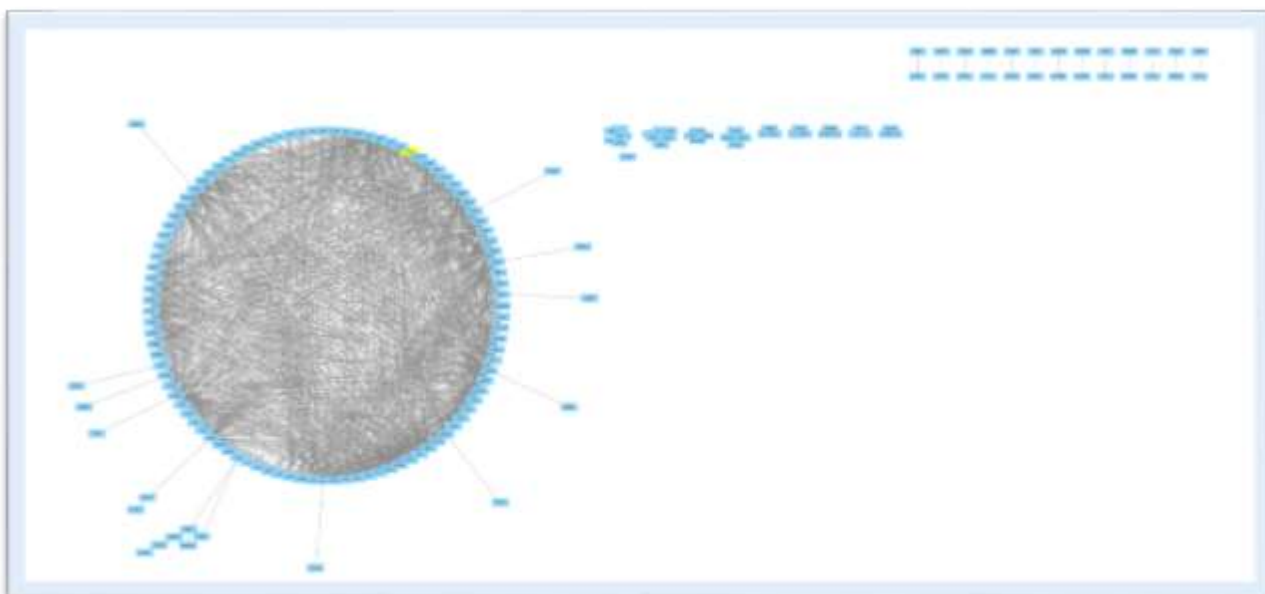
- 30 ژن در سینگالینگ PI3K

بهترین راه برای برقراری ارتباط بین داده ها و اطلاعاتی که تا به این مرحله کسب کرده ایم ، رسم شبکه است زیرا شبکه زبان سیستم های پیچیده است و به دلیل حالت گرافیکی ای که دارد تحلیلش راحت تر و قابل فهم تر است .

با استفاده از **starling** میتوانیم بر اساس اسامی ژن ها یک شبکه رسم و میانکشی های بین پروتئین ها را بررسی کنیم. در این شبکه 1037 ژن وجود دارد. **Sheet27**



در گام بعدی از نرم افزار cytoscapr استفاده می کنیم. این نرم افزار هم یک شبکه به ما می دهد که می توانیم ویژگی های ظاهری شبکه مانند رنگ و فاصله و نوع خطوط را به دلخواه تغییر دهیم. هر شبکه دو مولفه دارد: 1- nod -2 Edge
این نرم افزار قابلیت اضافه کردن اپلیکیشن های دیگری را جهت رسم و ارتقا کیفیت گرافیک شبکه ها دارد. sheet 28



دو مورد از پارامتر های مهم در شبکه ها :

Degree-1

betweenness centrality -2

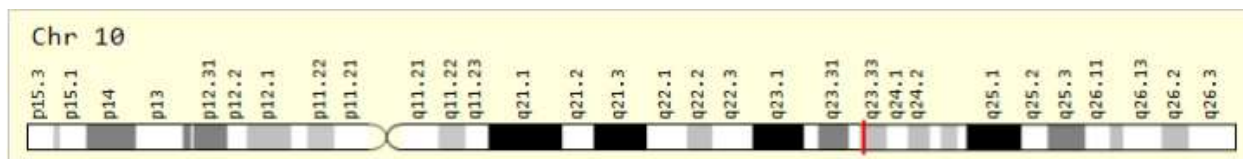
در بین تمام ژن هایی که تا به این جای کار بررسی شدند، 3 ژن هایی که بیشترین degree و BC را دارند را شناسایی می کنیم.

Gen	BC	Degree
KIF11	1/8271	61
CDK1	1/7477	65
NDC80	1/47	51

در نهایت از بین هزاران ژن که بررسی کردیم این سه ژن نقش کلیدی و مهم تری را در زمینه تحقیق‌ها داشتند. Sheet 29

با استفاده از سایت Gencards نقش و عملکرد هر یک از این پروتئین‌ها را بررسی و ارتباط دقیق‌تر را با موضوع که اختلاف بیان ژن تیروئید در دوران جنینی و بزرگسالی است، مورد مطالعه قرار می‌دهیم.

ژن **KIF11**:



Summaries for KIF11 Gene

[Entrez Gene Summary for KIF11 Gene](#)

This gene encodes a motor protein that belongs to the kinesin-like protein family. Members of this protein family are known to be involved in various kinds of spindle dynamics. The function of this gene product includes chromosome positioning, centrosome separation and establishing a bipolar spindle during cell mitosis. [provided by RefSeq, Jul 2008]

بر اساس اطلاعات بدست آمده از این سایت، عضوی از خانواده kenesin11 است که یک پروتئین حرکتی را کد می‌کند که عملکردهای مختلفی مانند موقعیت یابی کروموزومی، جداسازی سانتروزوم و ایجاد دوک قطبی در میتوز را برعهده دارد. از بیماری‌های مرتبط با این ژن می‌توان به عقب ماندگی ذهنی، Microcephaly اشاره کرد. این ژن در حضور هورمون تیروئید با گیرنده‌های تیروئید تعامل دارد.

Abstract

Purpose: Microcephaly with or without chorioretinopathy, lymphedema or intellectual disability (MCLID) is an autosomal dominant condition. Mutations in KIF11 have been found to be causative in approximately 75% of cases. This study describes the ocular phenotype in patients with confirmed KIF11 mutations.

PMID: 25996076

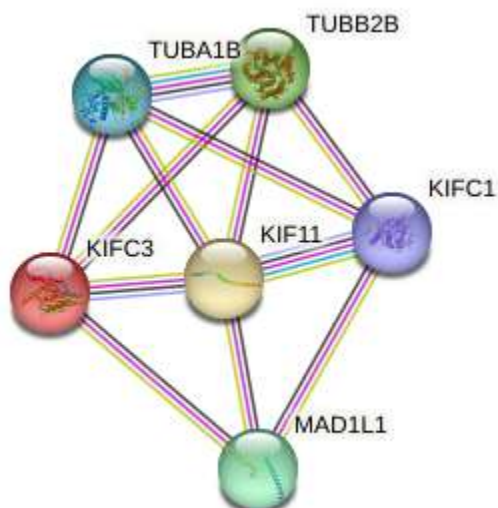
DOI: 10.1111/aos.12759

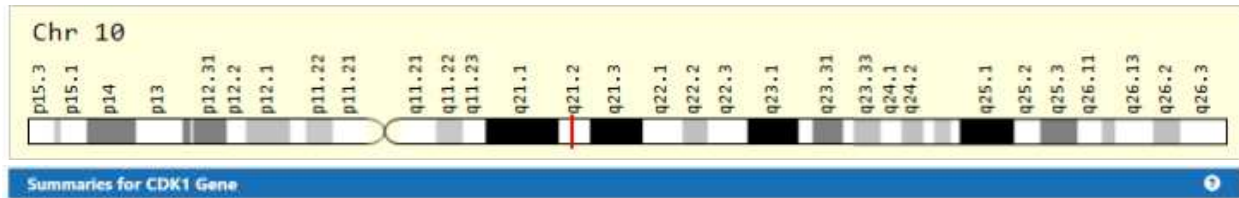
PMID: 31428438 PMCID: PMC6694292

DOI: 10.1038/hgv.2018.10

براساس این مقاله و توضیحاتی که از سایت gencard بدست آوردیم ، ژن KIF11 در ایجاد میکروسفالی تاثیر به سزایی دارد

جهش های KIF11 باعث ایجاد سندرم دیسپلازی autosomal - lymphedema غالب اتوزومی غالب، مرتبط با ناتوانی ذهنی نیز می شوند
ارتباط با سایر ژنها:





Summaries for CDK1 Gene

Entrez Gene Summary for CDK1 Gene

The protein encoded by this gene is a member of the Ser/Thr protein kinase family. This protein is a catalytic subunit of the highly conserved protein kinase complex known as M-phase promoting factor (MPF), which is essential for G1/S and G2/M phase transitions of eukaryotic cell cycle. Mitotic cyclins stably associate with this protein and function as regulatory subunits. The kinase activity of this protein is controlled by cyclin accumulation and destruction through the cell cycle. The phosphorylation and dephosphorylation of this protein also play important regulatory roles in cell cycle control. Alternatively spliced transcript variants encoding different isoforms have been found for this gene. [provided by RefSeq, Mar 2009]

این ژن عضوی از خانواده سرین تروئین کینازهاست و در چرخه سلولی اهمیت بالایی دارد
تغییر وضعیت هیستون ها، بیان ژن و حفظ هویت اپی ژنتیک سلول ها نیز قابل ذکر است.

Abstract

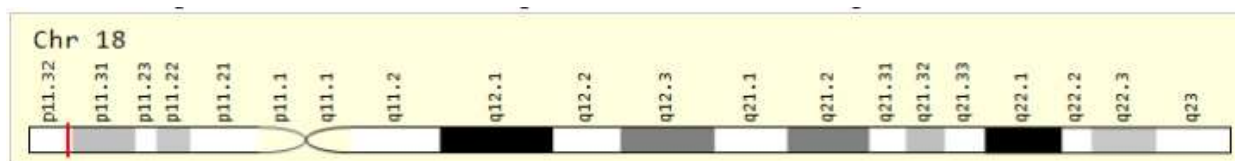
Cell proliferation exerts a high demand on protein synthesis, yet the mechanisms coupling the two processes are not fully understood. A kinase and phosphatase screen for activators of translation, based on the formation of stress granules in human cells, revealed cell cycle-associated kinases as major candidates. CDK1 was identified as a positive regulator of global translation, and cell synchronization experiments showed that this is an extramitotic function of CDK1. Different pathways including eIF2 α , 4EBP, and S6K1 signaling contribute to controlling global translation downstream of CDK1. Moreover, Ribo-Seq analysis uncovered that CDK1 exerts a particularly strong effect on the translation of 5'TOP mRNAs, which includes mRNAs encoding ribosomal proteins and several translation factors. This effect requires the 5'TOP mRNA-binding protein LARP1, concurrent to our finding that LARP1 phosphorylation is strongly dependent on CDK1. Thus, CDK1 provides a direct means to couple cell proliferation with biosynthesis of the translation machinery and the rate of protein synthesis.

PMID: 32040547

PMCID: PMC7054999

DOI: 10.1083/jcb.201906147

PMID: 32240602 PMCID: PMC7214218 DOI:
10.1016/j.molcel.2020.03.010



Summaries for NDC80 Gene

Entrez Gene Summary for NDC80 Gene

This gene encodes a component of the NDC80 kinetochore complex. The encoded protein consists of an N-terminal microtubule binding domain and a C-terminal coiled-coiled domain that interacts with other components of the complex. This protein functions to organize and stabilize microtubule-kinetochore interactions and is required for proper chromosome segregation. [provided by RefSeq, Oct 2011]

این ژن بخشی از کمپلکس NDC80 kinetochore را کد می‌کند.

این پروتئین برای سازماندهی و تثبیت تعاملات میکروتوبول - kinetochore و برای تفکیک مناسب کروموزوم مورد نیاز است

Abstract

Successful mitotic cell division is critically dependent on the formation of correct attachments between chromosomes and spindle microtubules. Microtubule attachments are mediated by kinetochores, which are large proteinaceous structures assembled on centromeric chromatin of mitotic chromosomes. These attachments must be sufficiently stable to transduce force; however, the strength of these attachments are also tightly regulated to ensure timely, error-free progression through mitosis. The highly conserved, kinetochore-associated NDC80 complex is a core component of the kinetochore-microtubule attachment machinery in eukaryotic cells. A small, disordered region within the Hec1 subunit of the NDC80 complex - the N-terminal "tail" domain - has been actively investigated during the last decade due to its roles in generating and regulating kinetochore-microtubule attachments. In this review, we discuss the role of the NDC80 complex, and specifically the Hec1 tail domain, at the kinetochore-microtubule interface, and how recent studies provide a more unified view of Hec1 tail domain function.

PMID: 32161753 PMCID: PMC7054225

DOI: 10.3389/fcell.2020.00043