تأثیر نتایج درمانی مرتبط با ویژگی‌های مولکولی کم‌خونی آپلاستیک کودکان[1]

فرزان رحمانی
مریم صادق‌آبادی
محمدرضا مزروعی

دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شریف،
گزارش‌کار پروژه درس مقدمه‌ای بر بیوانفورماتیک

18 بهمن 1403

چکیده

آنمی آپلاستیک شدید با مغز استخوان کم‌سلولی و سیتوپنی محیطی مشخص می‌شود. MSCها نقش مهمی در توسعه سلول‌های بنیادی خون‌ساز[2] و ایجاد یک محیط ریزمناسب برای خون‌سازی دارند. بررسی مولکولی مسیرهای نگهداری تلومر و پروفایل بیان ژنی MSCها می‌تواند برای مداخلات درمانی در بیماران کودکان مبتلا به آنمی آپلاستیک مهم باشد. همچنین با تحلیل نمونه‌های مشابه و خوشه‌بندی آن‌ها می‌توان بهتر نتایج به‌دست آمده را ارزیابی کرد. این مطالعه شامل ۱۰ بیمار کودک مبتلا به آنمی آپلاستیک و ۸ کودک سالم هم‌سن به عنوان گروه کنترل است. نمونه‌های خون محیطی از این افراد جمع‌آوری شد و طول تلومر لکوسیت‌ها با روش PCR زمان واقعی اندازه‌گیری شد. همچنین، داده‌های پروفایل بیان ژنی مبتنی بر ریزآرایه (GSE33812) از MSCهای پنج بیمار آنمی آپلاستیک کودکان، دانلود و تحلیل شد و با پنج فرد سالم به عنوان گروه کنترل مقایسه گردید. در نهایت با 5 دیتاست پروفایل بیان ژنی مشابه ادغام شده و ارزیابی از نتایج به‌دست آمده است. طول تلومر در بیماران کودک مبتلا به آنمی آپلاستیک به طور قابل‌توجهی کوتاه‌تر از افراد سالم هم‌سن بود. جالب اینکه، یک زیرگروه شامل دو بیمار کودک دارای طول تلومر متوسطی بودند که مشابه گروه کنترل بود. بر اساس تحلیل مسیر بیان ژنی مرتبط با نگهداری تلومر، ژن TERF2 به طور قابل‌توجهی در بیماران با تلومر کوتاه‌تر تنظیم منفی شده بود، اما در بیماران با تلومر متوسط تغییری مشاهده نشد. همچنین، پروفایل بیان ژنی MSCها نشان داد که سه ژن GAS2L3, MKI67, TMSB15A به طور افتراقی بیان شده و با نتایج درمانی مرتبط هستند. همچنین نتیجه کاهش ابعاد و خوشه‌بندی مجموعه داده ادغام نشان دهنده ارتباط نزدیک داده‌های کمکی و درستی نتایج به‌دست آمده درباره بیان ژن‌های معنا دار است و نتیجه می‌دهد که چه فرآیند سلولی در روند درمان بیماری موثر است. الگوی طول تلومر و پروفایل بیان ژنی MSCها و ادغام داده‌های نزدیک به‌هم و مسیرهای نگهداری آن‌ها ممکن است به عنوان یک نشانگر زیستی بالقوه برای پیش‌بینی نتایج درمانی آنمی آپلاستیک کودکان مورد استفاده قرار گیرد و به تولید دارو در درمان این بیماری کمک کند.

کلمات کلیدی

آنمی آپلاستیک، کودکان، تلومر، سلول‌های بنیادی مزانشیمی، درمان

---

[1] Implication of therapeutic outcomes associated with molecular characterization of paediatric aplastic anaemia
[2] HSCs

1- مقدمه

کم‌خونی آپلاستیک[3] یک بیماری نادر و خطرناک است که در اثر آسیب به مغز استخوان ایجاد می‌شود. مغز استخوان که وظیفه تولید سلول‌های خونی را بر عهده دارد، در این بیماری دچار کاهش شدید فعالیت شده و به‌جای سلول‌های خونی طبیعی، سلول‌های بسیار کمی تولید می‌کند یا به‌طور کامل عملکرد خود را از دست می‌دهد. این وضعیت باعث کاهش شدید سه نوع اصلی سلول‌های خونی می‌شود: گلبول‌های قرمز، گلبول‌های سفید و پلاکت‌ها. به این کاهش همزمان سلول‌ها اصطلاحاً پان‌سیتوپنی گفته می‌شود[1].

علائم بیماری کم‌خونی آپلاستیک عمدتاً ناشی از کاهش تعداد سلول‌های خونی است و شامل خستگی شدید به دلیل کمبود گلبول‌های قرمز و کاهش توانایی حمل اکسیژن، کبودی و خونریزی آسان مانند خونریزی لثه یا بینی به علت کاهش پلاکت‌ها، و عفونت‌های مکرر یا طولانی‌مدت به دلیل کاهش گلبول‌های سفید می‌شود، اگرچه عفونت‌ها در مراحل اولیه بیماری کمتر مشاهده می‌شوند. علاوه بر این، علائم عمومی مانند رنگ‌پریدگی، ضعف عضلانی، و کاهش تحمل فعالیت‌های بدنی نیز در میان بیماران رایج است.

برای بیماران مشکوک به کم‌خونی آپلاستیک، تشخیص سریع از اهمیت بالایی برخوردار است، زیرا در صورت عدم درمان، این بیماری می‌تواند تهدیدکننده حیات باشد. ابزارهای تشخیصی شامل آزمایش‌های خون (برای بررسی میزان سلول‌های خونی) و نمونه‌برداری از مغز استخوان برای ارزیابی سطح سلول‌سازی هستند.

برای درمان کم‌خونی آپلاستیک دو روش اصلی وجود دارد: درمان‌های سرکوب‌کننده سیستم ایمنی[4] و پیوند مغز استخوان[5]. درمان‌های سرکوب‌کننده سیستم ایمنی برای بیمارانی به کار می‌رود که امکان پیوند مغز استخوان ندارند و هدف آن کاهش پاسخ‌های خودایمنی بدن است که به سلول‌های مغز استخوان آسیب می‌رسانند. در مقابل، پیوند مغز استخوان، به‌ویژه در بیماران جوان‌تر و کودکانی که اهداکننده مناسبی دارند، روشی بسیار مؤثر است که عملکرد طبیعی مغز استخوان را بازمی‌گرداند. علاوه بر این، مراقبت‌های حمایتی، مانند تزریق خون و پیشگیری از عفونت‌ها، در کاهش عوارض بیماری نقش مهمی ایفا می‌کند[2].

شیوع کم‌خونی آپلاستیک در سراسر جهان متفاوت است. بر اساس داده‌های "مطالعه بین‌المللی آگرانولوسیتوز و کم‌خونی آپلاستیک"[6]، این بیماری در آسیا شایع‌تر است .عوامل محیطی، مانند تماس با مواد شیمیایی خاص، مصرف داروهای خاص، و عفونت‌های ویروسی، می‌توانند در بروز بیماری نقش داشته باشند. علاوه بر این، عوامل ژنتیکی و تنوع‌های قومی نیز ممکن است تأثیرگذار باشند[3].

عوامل محیطی مانند تماس با مواد شیمیایی و عفونت‌ها، به‌علاوه عوامل ژنتیکی، ممکن است خطر ابتلا به بیماری را افزایش دهند. کودکان مبتلا به AA، به دلیل وجود نقایص ارثی متعدد و حساسیت به درمان‌ها، نیاز به تشخیص و رویکردهای درمانی متفاوت دارند[4-7].

تلومرها، که به حفاظت از انتهای کروموزوم‌ها کمک می‌کنند، در AA اهمیت دارند. طول تلومر می‌تواند به‌عنوان یک شاخص تشخیصی و درمانی استفاده شود. مطالعات ژنومی، تفاوت‌های قابل‌توجهی در بیان ژن‌های دخیل در متابولیسم سلول و سیستم ایمنی میان بیماران و افراد سالم نشان داده‌اند[8-11].

سلول‌های بنیادی مزانشیمی[7] از مغز استخوان مشتق می‌شوند و نقش حیاتی در فرآیند خون‌سازی (هماتوپوییز) و تنظیم سیستم ایمنی ایفا می‌کنند. این سلول‌ها به‌عنوان یک ریزمحیط حمایتی برای سلول‌های بنیادی خون‌ساز عمل می‌کنند و با ترشح فاکتورهای رشد و سیگنال‌دهی، بقای این سلول‌ها را تضمین می‌کنند. در بیماران مبتلا به کم‌خونی آپلاستیک، اختلالات عملکردی قابل‌توجهی در MSCs مشاهده شده است، از جمله کاهش ظرفیت تکثیر سلولی، کاهش توانایی در حمایت از HSCs، و تغییرات در بیان ژن‌های مرتبط با چرخه سلولی و سیستم ایمنی[12-14].

مطالعات نشان داده‌اند که MSCs در بیماران AA توانایی خود را در تنظیم فرآیندهای زیستی مانند تقسیم سلولی، تکثیر، و تمایز از دست داده‌اند. این اختلالات می‌تواند منجر به نقص در بازسازی مغز استخوان و کاهش تولید سلول‌های خونی شود. همچنین، تغییر در بیان ژن‌های دخیل در سیگنال‌دهی ایمنی، مانند ژن‌های مرتبط با چسبندگی سلولی و برقراری ارتباط سلولی، نشان می‌دهد که MSCs نمی‌توانند محیط مناسبی برای تنظیم سیستم ایمنی فراهم کنند. این نقص‌ها می‌تواند باعث افزایش پاسخ‌های التهابی و آسیب به سلول‌های مغز استخوان شود[15-20].

علاوه بر این، مطالعات بیان کرده‌اند که MSCs در بیماران AA دارای پتانسیل کاهش‌یافته در ترمیم آسیب‌های سلولی و مقابله با استرس اکسیداتیو هستند. این موضوع نشان می‌دهد که عملکرد ضعیف MSCs ممکن است یکی از عوامل کلیدی در ایجاد و پیشرفت کم‌خونی آپلاستیک باشد. با توجه به این یافته‌ها، تحقیقات بیشتری برای درک بهتر تغییرات عملکردی MSCs و نقش آن‌ها در بیماری لازم است. همچنین، MSCs به‌عنوان یک هدف بالقوه برای مداخلات درمانی، مانند پیوند سلول‌های بنیادی یا تنظیم ژنتیکی، مورد توجه قرار گرفته‌اند[20-24].

---

[3] Aplastic Anemia
[4] IST
[5] BMT
[6] IAAAS
[7] MSCs

مطالعه حاضر به بررسی تاثیر مسیرهای نگهداری تلومر و وضعیت بیان ژن در MSCs می‌پردازد تا ارتباط آن‌ها با نتایج درمانی در کودکان مبتلا به AA تعیین شود.

2- داده‌ها و روش پیش‌پردازش

این بخش توضیحی جامع در مورد روش‌های به‌کاررفته در مطالعه ارائه می‌دهد. تحقیق بر روی 6 مجموعه‌داده که دارای بیماران مبتلا به کم‌خونی آپلاستیک و اهداکننده‌های سالم هستند، انجام شد. انتخاب بیماران مبتنی بر علائم کم‌خونی، کاهش سلول‌های خونی محیطی و مغز استخوان هیپوپلاستیک بوده است. افراد دارای سایر بیماری‌ها مانند HIV، هپاتیت، لنفوم، و بیماری‌های مزمن از مطالعه حذف شدند. نمونه‌های خون محیطی برای بررسی‌های مختلف جمع‌آوری شدند.

2-1- داده‌ها

مطالعه شامل سه گروه مختلف از بیماران بود: بیماران مبتلا به کم‌خونی آپلاستیک با طول تلومر طبیعی. بیماران مبتلا به کم‌خونی آپلاستیک با طول تلومر کاهش‌یافته. اهداکنندگان سالم.

DNA از نمونه‌های خون محیطی با استفاده از کیت‌های استاندارد استخراج شد. کیفیت و خلوص DNA با استفاده از اسپکتروفتومتر UV-Vis بررسی شد. برای اندازه‌گیری طول تلومر از واکنش زنجیره‌ای پلیمراز[8] استفاده شد. برای این منظور، پرایمرهای خاصی برای تلومر و ژن‌های منفرد مورد استفاده قرار گرفتند. نتایج طبق روش استاندارد تفسیر شدند. RNA کل از نمونه‌های خون محیطی استخراج شد و کیفیت آن با استفاده از ژل الکتروفورز تأیید شد. RNA استخراج‌شده برای سنتز cDNA با استفاده از ترانس‌کریپتاز معکوس مورد استفاده قرار گرفت. بیان ژن‌های منتخب مرتبط با نگهداری تلومر با استفاده از روش qPCR ارزیابی شد. بیان ژن‌ها نسبت به GAPDH به‌عنوان کنترل داخلی محاسبه و با اهداکنندگان سالم کالیبره شد[25-30].

بیان ژن سلول‌های بنیادی مزانشیمی با استفاده از داده‌های ریزآرایه موجود در پایگاه داده Gene Expression Omnibus[9] انجام شد. این تحلیل شامل 5 بیمار مبتلا به کم‌خونی آپلاستیک و 5 اهداکننده سالم بود. تفاوت‌های بیان ژن با استفاده از تغییرات برابر یا بیشتر از دو برابر[10] یا کمتر از نصف[11] گزارش شدند.

دیتاست GSE33812 شامل پروفایل‌های بیان ژنی سلول‌های بنیادی مزانشیمی استخراج‌شده از مغز استخوان پنج بیمار کودکان مبتلا به کم‌خونی آپلاستیک شدید[12] است. این مطالعه با استفاده از میکروآرایه‌های الیگونوکلئوتیدی انجام شده و هدف آن بررسی بیان سیتوکین‌ها و فاکتورهای رشد ضروری در MSCها بوده است. نتایج نشان داد که پانزده ژن مرتبط با تکثیر سلولی و سیتوکین‌ها به‌طور قابل‌توجهی در بیماران SAA کاهش یافته‌اند. این یافته‌ها با استفاده از واکنش زنجیره‌ای پلیمراز معکوس کمی[13] تأیید شدند.

به‌طور خاص، کاهش بیان ژن‌های CXCL12 (همچنین به‌نام SDF1) و فاکتور رشد هپاتوسیت[14] منجر به کاهش رشد سلولی و پاسخ به تمایز آدیپوژنیک، اما افزایش استئوژنز شد. بازگرداندن بیان ژن CXCL12 یا افزودن HGF به‌صورت برون‌زا، عملکرد سلولی MSCهای بیماران SAA را تا حدی بازیابی کرد. این نتایج نشان می‌دهد که بیان ژن در MSCهای بیماران SAA به‌طور کلی کاهش یافته و ژن‌های CXCL12 و HGF نقش‌های حیاتی در تنظیم رشد و تمایز سلولی ایفا می‌کنند.

دیتاست GSE165870 شامل داده‌های توالی‌یابی RNA[15] از سلول‌های بنیادی و پیش‌ساز خونساز Lin-CD34 (HSPCs)+ استخراج‌شده از مغز استخوان سه اهداکننده سالم و بیماران مبتلا به کم‌خونی آپلاستیک درمان‌نشده است. برای تهیه این نمونه‌ها، ابتدا سلول‌های تک‌هسته‌ای از آسپیراسیون مغز استخوان با استفاده از جداسازی گرادیان چگالی فیکول جدا و در ۹۰٪ FBS و ۱۰٪ DMSO در نیتروژن مایع نگهداری شدند. پس از ذوب‌کردن، سلول‌ها با آنتی‌بادی‌های انسانی شامل کوکتل لاینئیج (CD3، CD14، CD16، CD19، CD20، CD56) و CD34 رنگ‌آمیزی شدند. سپس، سلول‌های Lin-CD34+ HSPCs با استفاده از دستگاه BD FACSAria III جداسازی و مستقیماً در بافر لیز سلولی قرار گرفتند. مراحل رونویسی معکوس و تعویض الگو بر اساس پروتکل Smart-seq2 با برخی تغییرات انجام شد. توالی‌یابی با استفاده از پلتفرم Illumina NovaSeq 6000 و تولید خوانش‌های ۱۵۰ جفت‌باز انجام شد.

در پردازش داده‌ها، توالی‌های پرایمر TSO، دنباله پلی‌آ، آداپتور و توالی‌های با کیفیت پایین حذف شدند. توالی‌های تمیز به ژنوم انسانی GRCh38 با استفاده از نرم‌افزار STAR هم‌تراز شدند و خوانش‌های تکراری با استفاده از Picard حذف شدند. برای تحلیل‌های بیشتر، از نرم‌افزارهای MATS و DaPars به ترتیب برای بررسی اتصالات جایگزین و پلی‌آدنیلاسیون جایگزین استفاده شد. کمی‌سازی ژن‌ها با استفاده از HTSeq انجام شد. این دیتاست به محققان امکان می‌دهد تا تفاوت‌های بیان ژنی بین HSPCs افراد سالم و بیماران AA را بررسی کنند و به درک بهتری از مکانیسم‌های مولکولی مرتبط با AA دست یابند.

---

[8] qPCR
[9] GEO
[10] upregulation
[11] downregulation
[12] SAA

[13] qRT-PCR
[14] HGF
[15] RNA-seq

دیتاست GSE28974 شامل پروفایل‌های بیان ژنی سلول‌های بنیادی مزانشیمی انسان است که با استفاده از پلتفرم Illumina BeadChips Human-6v2.0 مورد بررسی قرار گرفته‌اند. در این مطالعه، ۱۲ نمونه MSC مورد آنالیز قرار گرفته است.

این دیتاست بخشی از یک سوپر سری بزرگ‌تر است که شامل GSE33383 و GSE42352 می‌باشد و به شناسایی ژن‌های محرک استئوسارکوما از طریق تحلیل یکپارچه داده‌های تعداد نسخه و بیان ژن می‌پردازد. محققان می‌توانند از این داده‌ها برای درک بهتر مکانیسم‌های مولکولی مرتبط با MSCها و نقش آن‌ها در بیماری‌هایی مانند استئوسارکوما استفاده کنند.

دیتاست GSE3807 شامل پروفایل‌های بیان ژنی سلول‌های T CD3+ استخراج‌شده از مغز استخوان و خون محیطی بیماران مبتلا به کم‌خونی آپلاستیک و داوطلبان سالم است. در این مطالعه، از ریزتراشه‌های Affymetrix HG_U133A برای تحلیل بیان ژن‌ها استفاده شده است. نمونه‌ها شامل چهار بیمار AA (سه زن و یک مرد، ۱۹ تا ۷۰ ساله) بودند که از هر کدام نمونه‌های مغز استخوان و خون محیطی جمع‌آوری و RNA کل استخراج و ترکیب شد. همچنین، دو بیمار دیگر (یک مرد ۶۴ ساله و یک زن ۷۰ ساله) مورد بررسی قرار گرفتند که از آن‌ها نمونه‌های قبل و بعد از درمان جمع‌آوری شد. نمونه‌های کنترل سالم نیز از سه اهداکننده برای خون محیطی و دو اهداکننده برای مغز استخوان تهیه شدند.

هدف این مطالعه، مقایسه بیان ژنی در سلول‌های T بیماران AA و افراد سالم، و همچنین بررسی تغییرات بیان ژنی قبل و بعد از درمان در بیماران AA بود. این داده‌ها می‌توانند به درک بهتر مکانیسم‌های مولکولی مرتبط با AA و پاسخ به درمان کمک کنند. نتایج این تحقیق در مقاله‌ای با شناسه PubMed 17052335 منتشر شده است.

دیتاست GSE29105 به بررسی پروفایل‌های بیان ژنی سلول‌های بنیادی خون‌ساز انسانی پرداخته است که از خون بند ناف استخراج شده‌اند. در این مطالعه، سلول‌ها بر اساس نشانگرهای سطحی CD34، CD38، CD45RA، Thy1 و CD49f جداسازی شده‌اند تا جمعیت‌های مختلف HSCs و پیش‌سازهای چندتوان را شناسایی کنند. هدف اصلی این تحقیق، شناسایی ژن‌هایی است که در HSCها بیشتر بیان می‌شوند و ممکن است با عملکرد سلول‌های بنیادی و خودنوسازی آن‌ها مرتبط باشند.

برای انجام این آنالیز، RNA کل از جمعیت‌های سلولی جداسازی‌شده با استفاده از ریزتراشه‌های Illumina HumanHT-12 V3.0 expression beadchip مورد بررسی قرار گرفته است. این داده‌ها می‌توانند به درک بهتر مکانیسم‌های مولکولی مرتبط با عملکرد HSCها کمک کنند و ژن‌های کلیدی مرتبط با خودنوسازی و تمایز این سلول‌ها را شناسایی نمایند. نتایج این مطالعه در مقاله‌ای با شناسه PubMed 21737740 منتشر شده است.

دیتاست GSE113033 به بررسی پروفایل بیان ژنی زیرجمعیت‌های جدید سلول‌های بنیادی خون‌ساز (HSC) تعریف‌شده توسط بیان CD35 و جمعیت‌های پیش‌ساز در خون‌سازی بالغ انسان می‌پردازد. در این مطالعه، جمعیت‌های هدف با استفاده از دستگاه FACS Aria III جداسازی شدند و RNA استخراج‌شده از این جمعیت‌ها با استفاده از میکروارای Agilent مورد آنالیز قرار گرفت. نتایج نشان داد که HSCهای CD35+ در مقایسه با HSCهای CD35- و جمعیت‌های پیش‌ساز، بیان کمتری از ژن‌های مرتبط با چرخه سلولی یا تمایز دارند که نشان‌دهنده وضعیت خفته یا نابالغ آن‌ها است.

این دیتاست شامل ۱۸ نمونه است که در پلتفرم Agilent-072363 SurePrint G3 Human GE v3 8x60K Microarray 039474 مورد بررسی قرار گرفته‌اند. نمونه‌ها شامل تکرارهای بیولوژیکی از HSCهای CD35+، HSCهای CD35-، سلول‌های پیش‌ساز چندتوانی[16]، پیش‌سازهای میلوئیدی مشترک[17]، پیش‌سازهای گرانولوسیت-ماکروفاژ[18] و پیش‌سازهای مگاکاریوسیت-اریتروایِد[19] هستند. این مطالعه توسط محققان دانشگاه کیوشو در ژاپن انجام شده است.

2-2- روش پیش‌پردازش

با توجه به استفاده از مجموعه‌داده های متفاوت که هر کدام تکنولوژی مختلفی دارند و همچنین پردازش داده های مجموعه داده مقاله به شکل خام[20] روش‌های پیش پردازش گوناگونی را انجام دادیم که به شرح زیر هستند:

2-2-1- داده‌های موجود در مقاله اصلی

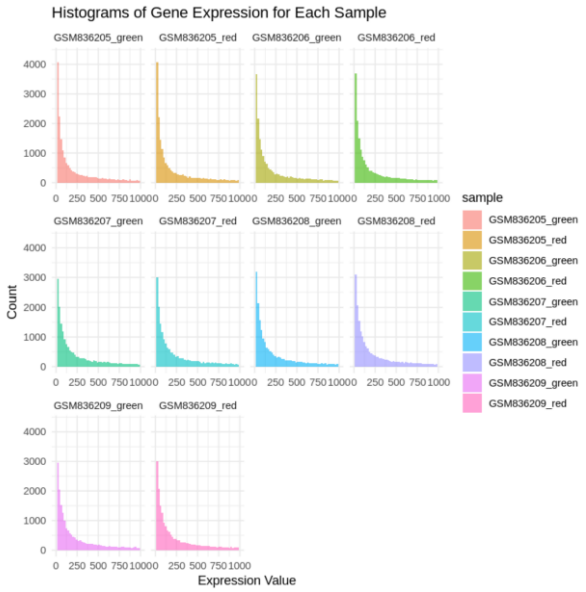
مجموعه داده خام با شماره دسترسی "GSE33812" از پایگاه داده GEO دانلود شد. داده های استخراج شده از هر نمونه با هم ترکیب شدند و شناسه های نمونه ها تحت عنوان "sampleID" جهت تسهیل در مراحل پیش پردازش، به داده ها اضافه شد.

هر نمونه به صورت دو کاناله ثبت شده است که شامل مقادیر سیگنال برای کانال های سبز(g) و قرمز (r) میباشد. در این نوع ریزآرایه ها دو رنگ فلورسنت مختلف برای اندازه گیری بیان ژن استفاده میشود.کانال سبز معمولا برای نمونه های شاهد یا کنترلی و کانال قرمز معمولا برای نمونه های بیمار استفاده میشود.

---

[16] MPP
[17] CMP
[18] GMP
[19] MEP
[20] raw

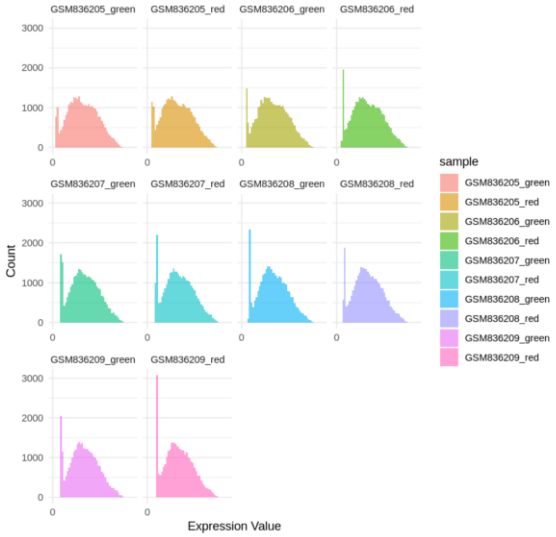
داده ها برای شناسایی ردیف های تکراری بررسی شدند که شامل پروب هایی با مقادیر گمشده برای تمامی ویژگی ها بود.تمامی ردیف های تکراری که ممکن بود ناشی از خطای پردازش باشند،حذف شد.
پروب هایی که برای کنترل کیفیت آزمایش استفاده میشدند،حذف شد و تحلیل ها تنها براساس پروب های مربوط به ژن های هدف انجام شد.
داده‌ها از نظر وجود مقادیر گمشده بررسی شدند.ستون "Locuslink" که به طور کامل فاقد داده بود حذف شد.
در مرحله بعدی برای ساخت ماتریس بیان ژن،داده‌ها بر اساس ترکیب نام ژن و شناسه نمونه گروه‌بندی شدند. ژن‌هایی که در یک نمونه بیش از یک بار تکرار شده بودند شناسایی شدند. برای ژن‌های تکراری، میانگین بریده [21](با حذف ۱۰٪ داده‌های انتهایی) از سیگنال محاسبه شدند و برای ژن های غیرتکراری مقادیر اصلی سیگنال حفظ شد.
در نهایت، ماتریس بیان ژن نهایی ایجاد شد که ژن‌ها به عنوان سطرها و نمونه‌ها با کانال رنگی جداگانه به عنوان ستون‌ها قرار گرفتند.



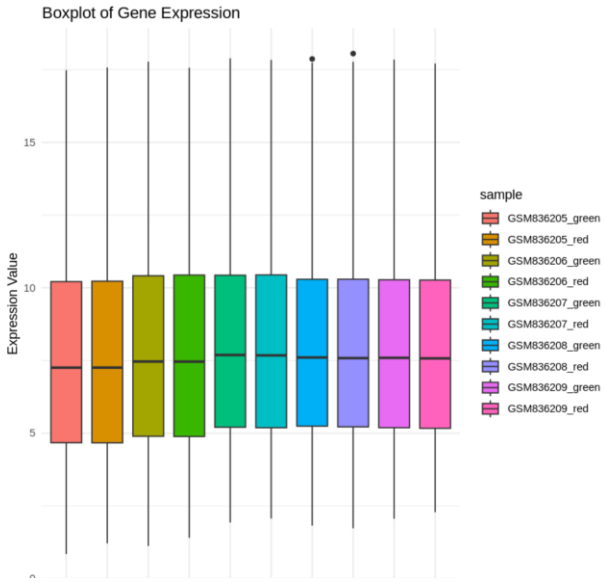
شکل 1.نمودار هیستوگرام بیان ژن ها در هر یک از 10 نمونه پیش از پردازش

خلاصه آماری داده ها نشان دهنده تنوع گسترده ای در بیان ژن ها بین نمونه های مختلف بودند. همچنین انحراف به راست هیستوگرام توزیع بیان ژن ها برای هر نمونه در شکل 1 نشان میدهد بیشتر داده‌ها در بازه‌های پایین‌تر از ۱۰۰۰ قرار داشتند، اما تعدادی از ژن‌ها دارای مقادیر بسیار بالایی بودند که نشان‌دهنده بیان قوی یا آرتیفکت‌های فنی هستند که لزوم نرمالسازی را برجسته میکند.
با توجه به دریافت داده خام، اصلاح پس زمینه انجام شد.سپس نرمالسازی با استفاده از روش "Quantile Normalization" انجام شد که به همسان‌سازی توزیع سیگنال‌ها بین نمونه‌های مختلف کمک می‌کند.

همچنین با توجه به مقادیر خلاصه آماری، داده های نرمال شده با استفاده از تبدیل لگاریتمی تغییر یافتند که به کاهش واریانس و نرمالسازی بهتر داده ها کمک میکند.
پس از تبدیل لگاریتمی، یکنواختی توزیع مقادیر بیان ژن میان نمونه‌ها نشان می‌دهد که داده‌ها برای انجام تحلیل‌های آماری و بررسی تغییرات بیان ژنی میان گروه‌های مختلف، قابل اعتماد هستند.
نمودار جعبه‌ای در شکل 3 نشان می‌دهد که نمونه‌های مختلف دارای توزیع مشابهی از مقادیر بیان ژن هستند، که این امر بیانگر نرمال‌سازی موفق داده‌ها و کاهش تفاوت‌های تکنیکی میان نمونه‌هاست. علاوه بر این، بررسی هیستوگرام‌های بیان ژن برای هر نمونه در شکل 2 نشان می‌دهد که اکثر داده‌ها در محدوده‌های مشابهی نزدیک به توزیع نرمال قرار دارند و پراکندگی کلی نمونه‌ها همگن است.



شکل 2.نمودار هیستوگرام بیان ژن ها در هر یک از 10 نمونه پس از پردازش



شکل 3.توزیع داده های 10 نمونه پس از پردازش

[21] Trimmed Mean

داده‌ها با موفقیت پیش‌پردازش شدند و آماده برای تحلیل‌های پیشرفته‌تر مانند شناسایی ژن‌های دیفرانسیل بیان شده، تحلیل مسیرهای زیستی و بررسی اثرات بیولوژیکی مرتبط با آنمی آپلاستیک می‌باشد.

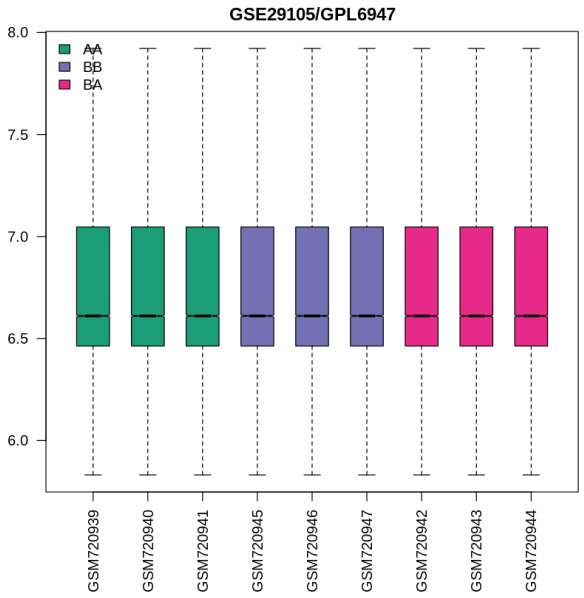
2-2-2- افتراق بیان ژنی

این مرحله شامل تطبیق مقادیر بیان ژنی با مدل خطی جهت تخمین اثرات هر شرایط بر بیان ژن‌های مختلف است. با استفاده از این مدل، می‌توان تأثیرات سیستماتیک ناشی از شرایط مختلف را بر داده‌ها بررسی کرد و میزان اختلاف‌های بیان ژنی را به‌صورت کمی ارزیابی نمود. سپس روش تعدیل بیزی تجربی، اطلاعات مربوط به تمامی ژن‌ها را ترکیب کرده و تخمین‌های معتبرتری از تغییرات بیان ارائه می‌دهد.ژن‌های با بیان دیفرانسیل شناسایی شدند که تفاوت معنی‌داری بین شرایط موردمطالعه نشان میدهند که مقادیر LogFold و p.value را برای هر ژن محاسبه کرده است.

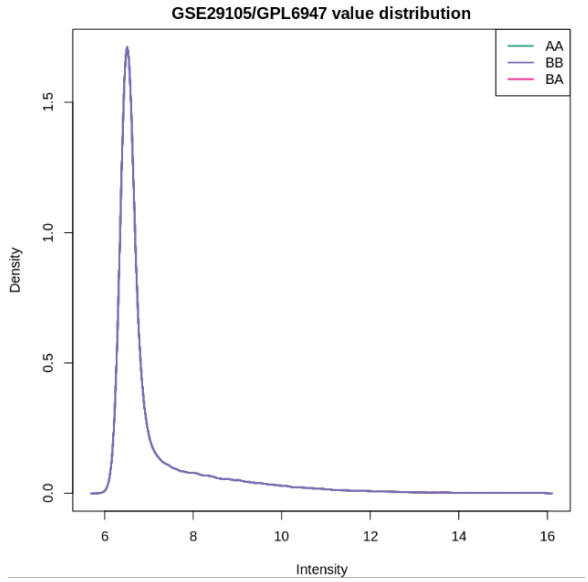
2-2-3- داده‌های افزوده مشابه مقاله اصلی به شکل مستقل از هم

کد های کامل پیش‌پردازش و تحلیل این مجموعه داده ها بـه صـورت مجزا در نوتبوک های ژوپیتر ضمیمه شده و با کامنت های شفاف کننده آمده اند. آنهـا را مـی تـوان در فایـل هـای GSE28974.ipynb، GSE3807.ipynb، GSE29105.ipynb، GSE113033.ipynb و GSE165870.ipyn مشاهده کـرد. در اینجا به طور مختصر آنهـا را توضـیح داده و سـپس خروجـی هـای موجـود در یکـی از آنهـا را میبینـیم. بـه طـور خـاص مجموعـه داده GSE29105 را در اینجـا بررسـی میکنـیم و توجـه داریـم کـه 4 مجموعه داده دیگر نیز مشابه همین مجموعه داده هستند.

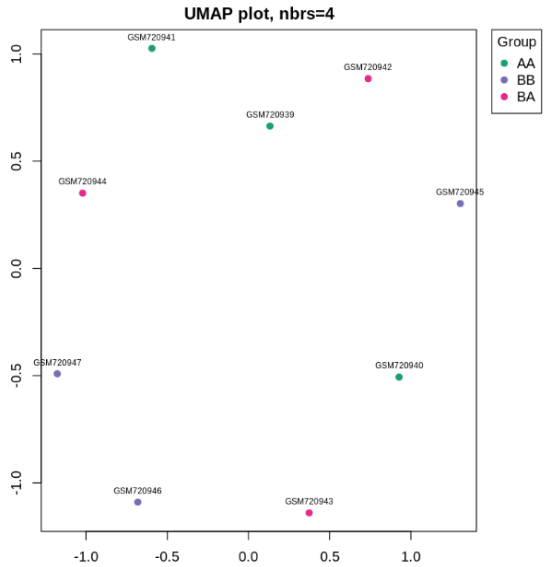
- در کد R موجود مشاهده می شود که ابتدا بـه نصـب package هـای ضـروری نظیـر DESeq2، GEOquery و غیـره مـی پردازیم. سپس کتابخانه های ضروری را import میکنیم.
- در مرحله بعد داده ها را از پایگاه داده GEO دانلود میکنیم. سپس نـام ستون ها را با استفاده از فایل annotation اصلاح میکنیم.
- حال به گروه بندی نمونه ها با توجه به توضیحات خاص مجموعه داده میپردازیم.
- حال که ماتریس بیان ژن ها در نمونه ها را داریم، به نرمال سازی داده ها با اسـتفاده از روش quantile normalization مـی پردازیم. سپس log2 transformation را انجام مـی دهـیم. در نهایت به مدیریت missing values و حذف یا پـر کـردن آنها با استفاده از روش مناسب نظیر مقدار میانه می پردازیم.
- در مرحله بعد نمودار جعبه ای را رسم میکنیم تـا از نرمـال سـازی درست مطمئن شویم که در شکل زیر مشاهده میکنید. همان طور که میبینید داده ها به خوبی نرمال شده اند.
- در نهایت نیز، با رسم نمودار توزیع مقـادیر و نمونـه هـای کـاهش ابعاد یافته نمونه هـا بـا الگـوریتم unmap از درسـتی کـد خـود اطمینان حاصل میکنیم.



شکل 4.نمودار جعبه ای پس از نرمال سازی



شکل 5.نمودار توزیع مقادیر پس از پیش پردازش

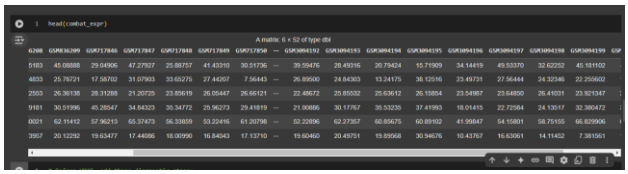


شکل 6.نمودار کاهش ابعاد نمونه ها بعد از پیش پردازش

2-2-4- داده‌های افزوده مشابه مقاله اصلی به شکل ترکیب شده[22]

کد های کامل پیش‌پردازش و تحلیل این مجموعه داده ها بـه صـورت ترکیب شده و یکجا در نوتبوک ژوپیتر ضمیمه شده و بـا کامنـت هـای شـفاف کننـده آمـده اسـت. آن را مـی تـوان در فایـل bio_proj_merged_5_datasets.ipynb مشاهده کرد. توجه داریم که 5 مجموعه داده را با هـم ترکیـب کـردیم کـه یکـی از آنهـا مجموعه داده اصلی مقاله خواسته شـده اسـت. 4 مجموعـه داده دیگـر مجموعه داده های مشابه آن هستند. همچنین برای این قسـمت چـون تکنولوژی مجموعه داده GSE165870 از نوع RNA-seq data است از آن صرف نظر کردیم چـرا کـه سـایر مجموعـه داده هـا دارای تکنولوژی Microarray هسـتند. در اینجـا بـه طـور مختصـر آن را توضیح داده و سپس خروجی های موجود در آن را میبینیم. توجه داریـم چالش اصلی در این قسمت این است که داده هـای هـر مجموعـه داده دارای فناوری و platform های گونـاگونی هسـتند. ایـن منجـر بـه تفاوت نامگذاری سـطر هـای مـاتریس بیـان ژن بـرای نـام ژن هـا و همچنین ژن های غیر مشترک موجود می شود.

- در کد R موجود مشاهده می شود که ابتدا بـه نصـب package هـای ضـروری نظیـر limma، GEOquery و غیـره مـی پردازیم. سپس کتابخانه های ضروری را import میکنیم.
- با توجه به چالش فناوری های مختلف هر مجموعه داده یک تـابع با نام process_dataset تعریف کردیم که بعـد از دانلـود هـر مجموعه داده با توجه به فناوری مجموعـه داده و annotation های موجود همه مجموعه داده ها را بـه یـک فضـای نامگـذاری نگاشت کند.این تابع پیچیده ای است که توضـیح جزییـات آن در اینجا نمی گنجد ولی با مراجعه به نوتبوک ضمیمه شده و خوانـدن کامت ها می توانید جزییات آن را مشاهده کنید.
- سـپس، بـا تـابع process_dataset ایـن 5 مجموعـه داده را دانلود و پردازش میکنیم. همچنین پس از اینکه نـام ژن هـا یکـی شد با اشتراک گرفتن بین 5 مجموعه داده، ژن های مشـترک آن ها را پیدا میکنیم.
- در مرحله بعد، یک ماتریس که حاصل ترکیب این 5 مجموعه داده است می سازیم که سطر های آن ژن های مشترک و سـتون هـا آن نمونه های مختلف در مجموعه داده ها هستند.
- حال با کتابخانه ComBat عملیـات Batch correction را انجام می دهیم که بسیار ضروری است چرا که مجموعـه داده هـا فناوری های گوناگونی دارند.
- در نهایـت نیـز مقـادیر NA/NaN را از مـاتریس خـود حـذف میکنیم. در شکل زیـر میتوانیـد 6 سـطر اول مـاتریس حاصـل از ترکیب مجموعه داده ها مشاهده کنید.



شکل 7.مجموعه داده ها پس از ترکیب و پیش پردازش

3- نتایج

همون طور که در بخش قبل توضیح دادیم مـا بـه سـه شـیوه مختلـف مجموعه داده ها را ارزیابی کردیم. لذا نتایج هر قسمت و تحلیـل هـای آنها را در ادامه و به شکل تفکیک شده مشاهده میکنیم.

3-1-1- داده‌های موجود در مقاله اصلی

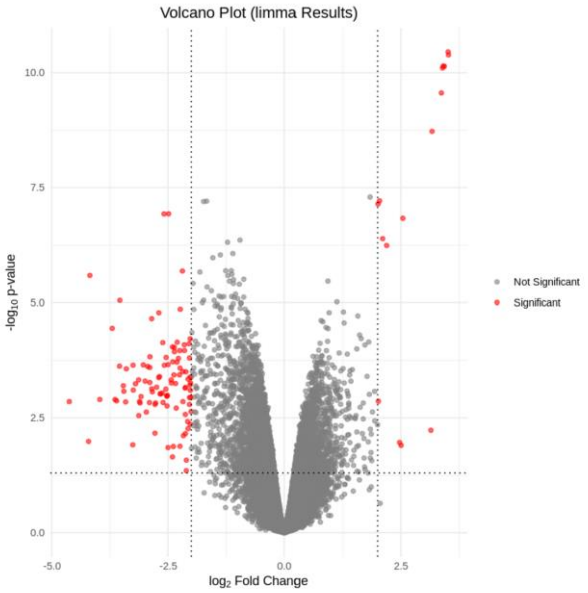
در تحلیل داده‌های زیستی، تعیین آستانه‌های معناداری آماری برای تشخیص تغییرات مهم در بیان ژن ضروری است.طبق مقاله، در این فرآیند دو معیار کلیدی به کار می‌رود $p.value<0.05$ که نشان‌دهنده‌ی احتمال بسیار کم در انتخاب تصادفی ژن معنادار است و $|logFC| \geq 2$ که نشان‌دهنده‌ی تغییرات قابل‌توجه در سطح بیان ژن است.

نمودار آتشفشانی[23] در شکل 8 که برای نمایش توزیع ژن‌های دیفرانسیل بیان شده بر اساس مقادیر p.value و logFC استفاده میشود.این نمودار نشان می‌دهد که از بین 30598، 119 ژن به عنوان با رنگ قرمز شناسایی شده‌اند که ممکن است در فرآیندهای زیستی مرتبط با بیماری نقش داشته باشند. این ژن‌ها به دلیل تغییرات بزرگ در بیان و معناداری بالا، به‌عنوان کاندیدهای بالقوه برای مطالعات بیشتر پیشنهاد می‌شوند.

---

[22] merged

[23] Volcano plot



شکل 8.نمودار آتشفشانی برای نمایش نتایج تحلیل بیان تفاضلی ژن‌ها با استفاده از پکیج limma

در نمودار heatmap در شکل 9، رنگ ها بیانگر میزان بیان ژن ها در هر نمونه هستند. همانطور که مشاهده میشود که بیان ژن های معنی دار به طور محسوسی بین دو گروه سالم و بیمار متفاوت است، به ویژه در ژن های کلیدی آپلاستیتیک. این تفاوت‌ها در الگوی بیان ژنی نمایانگر وجود مسیرهای زیستی خاصی است که در شرایط بیماری و نرمال تغییر می‌کنند.

شکل 9.نمودار شباهت نمونه ها بر اساس ژن های انتخاب شده

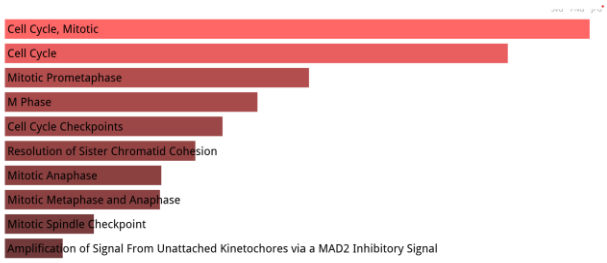
با بررسی ژن های معنادار سه ژن AGT،MK16 وGAS2L3 به مشاهده میشود که در مقاله مرجع نیز به آن اشاره شده بود. ژن AGT در تمایز سلول‌های بنیادی مزانشیمی (MSCs) به سلول‌های چربی نقش دارد. خاموش شدن این ژن ممکن است به بروز اختلالات ایمنی منجر شود.

ژن GAS2L3 که با ساختار اسکلت سلولی مرتبط است. مطالعات رو مدل حیوانی نشان داده حذف این ژن منجر به کاهش تکثیر سلول‌های قلبی و هیپرتروفی سلولی می‌شود که می‌تواند به عنوان یک نشانگر تکثیری در کم‌خونی آپلاستیک نیز مطرح باشد.

همچنین، ژن MK167 که در تکثیر سلولی نقش دارد، به عنوان یک نشانگر تکثیر در بیماران مبتلا به کم‌خونی آپلاستیک در نظر گرفته شده است.

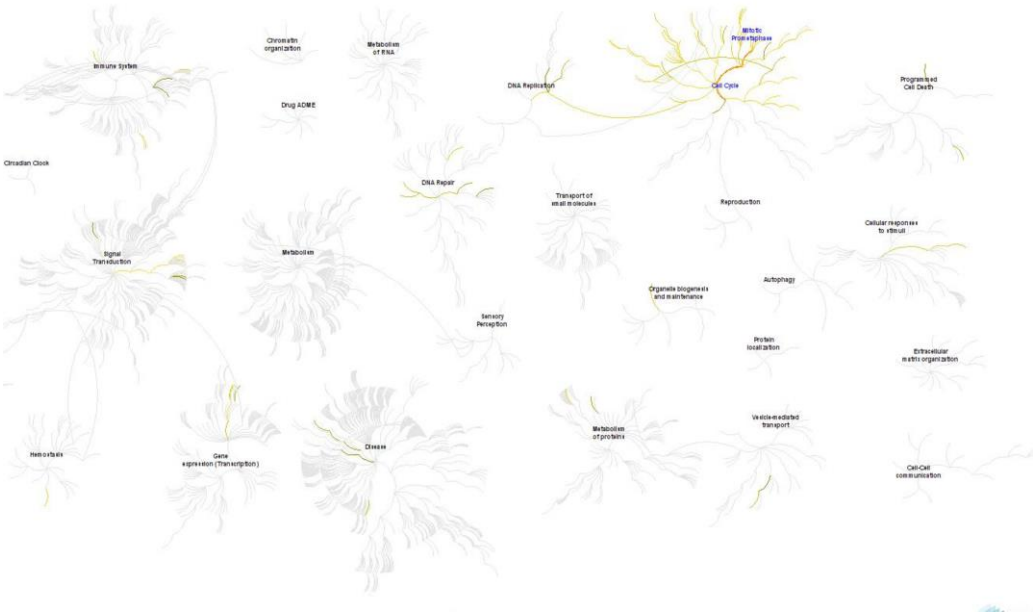
در ادامه به بررسی اهمیت مسیرهای زیستی مرتبط با سلول‌های مزانشیمی در بیماران مبتلا به کم‌خونی آپلاستیک می‌پردازیم. برای این تحلیل از پایگاه داده Enrichr استفاده شده است. همان‌طور که در شکل 10مشاهده می‌شود، مسیرهای زیستی مرتبط با چرخه سلولی و تقسیم میتوزی در این سلول‌ها دچار تغییرات قابل توجهی شده‌اند. در صدر این مسیرها، چرخه سلولی میتوزی [24] و تنظیم کلی چرخه سلولی [25] قرار دارند که بیانگر تغییرات گسترده در فرآیندهای تقسیم سلولی است. همچنین مسیرهای پروفازمیتوزی، فاز M و نقاط بازرسی چرخه سلولی [26] نشان می‌دهند که ناهنجاری‌هایی در کنترل تقسیم سلولی و اطمینان از صحت انتقال کروموزوم‌ها رخ داده است.اختلال در این مسیر می‌تواند منجر به توقف چرخه‌ی سلولی، تکثیر غیرقابل‌کنترل سلول‌ها یا مرگ سلولی برنامه‌ریزی‌شده (آپوپتوز) شود.

این یافته‌ها نشان می‌دهند که اختلال در تقسیم سلولی و تنظیم چرخه سلولی یکی از مکانیسم‌های کلیدی در سلول‌های مزانشیمی بیماران آپلاستیک است. چنین تغییراتی ممکن است بر تکثیر و پایداری این سلول‌ها اثر گذاشته و در کاهش توان بازسازی بافتی و خون‌سازی مؤثر باشد.



شکل 10.مسیرهای زیستی حاصل از تغییرات بیان ژنی در بیماران مبتلا به کم‌خونی آپلاستیک

[24] Cell Cycle, Mitotic
[25] Cell Cycle
[26] Cell Cycle Checkpoints



شکل 11.نمودار اهمیت ژن های انتخاب شده در فعالیت های زیستی بدن با تمرکز برفرآیندهای چرخه سلولی، تنظیم ایمنی و ترمیم DNA

بخش‌های زرد در شکل 11 که از منبع Reactom[40] به دست آمده است، نشان‌دهنده‌ی مسیرهای زیستی تحت تأثیر ژن‌های مشخص‌شده در سلول‌های مزانشیمی آپلاستیک هستند.

همان‌طور که در نمودار مشاهده می‌شود، ، مسیر چرخه‌ی سلولی یکی از مهم‌ترین مسیرهای تحت تأثیر است که شامل مراحل کلیدی مانند پرو متافاز میتوزی می‌شود.یکی از مسیرهای کلیدی شناسایی شده، در ترمیم DNA نقش مهمی دارد در حفظ یکپارچگی ژنوم و جلوگیری از تغییرات مخرب در سلول‌های مزانشیمی آپلاستیک دارد.

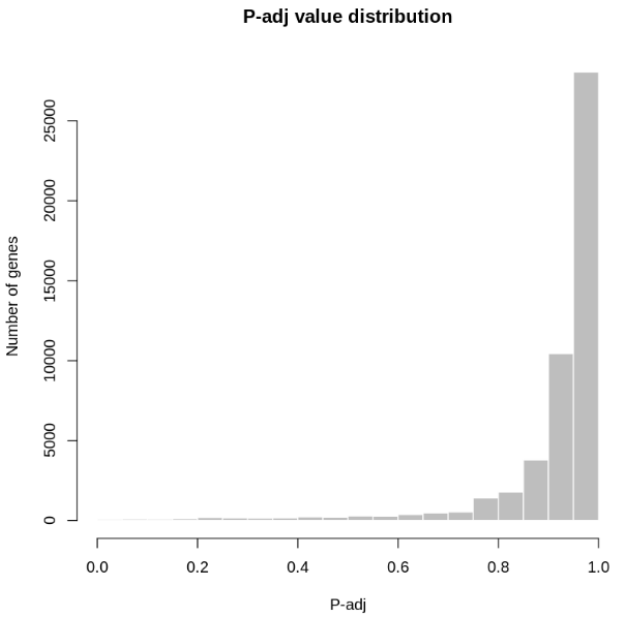
همچنین مسیرهای مرتبط با سیستم ایمنی نیز مشاهده می‌شوند که نشان‌دهنده تأثیرات گسترده این تغییرات ژنی بر روی فرآیندهای تنظیمی و پاسخ‌های ایمنی است. این مسیرها ممکن است با فرآیندهای التهابی مرتبط باشند که در آنمی آپلاستیک نقش کلیدی ایفا می‌کنند.

این یافته‌ها تأیید می‌کنند که ناهنجاری‌های موجود در این مسیرها ممکن است به نقص در بازسازی سلول‌های بنیادی و اختلال در عملکرد طبیعی سلول‌های مزانشیمی منجر شوند.

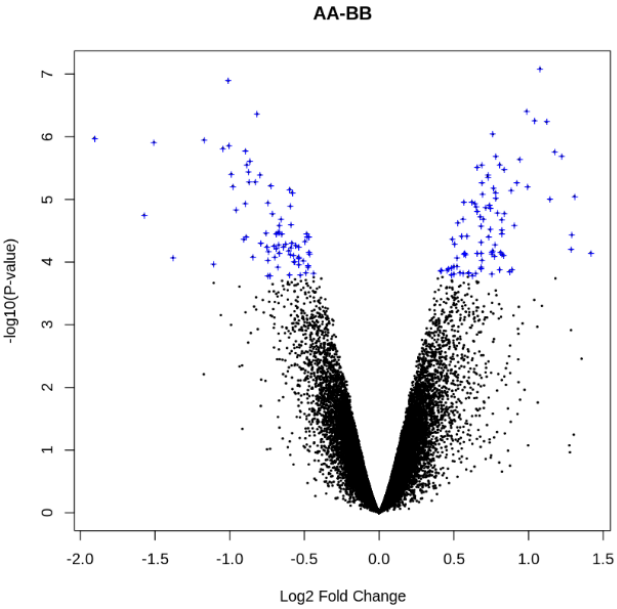
2-1-3- داده‌های افزوده مشابه مقاله اصلی به شکل مستقل از هم

نمودار و خروجی های به دست آمده به شکل کامل در فایل های GSE28974.ipynb، GSE3807.ipynb، GSE29105.ipynb، GSE113033.ipynb و GSE165870.ipynb قابل مشاهده هستند. در اینجا به توضیح مختصری از یکی از آنها می پردازیم چرا که شبیه به هم بوده و همچنین در صورت توضیح هر 5 مجموعه داده به شکل جدا جدا گزارش طولانی خواهد شد و انسجام خود را از دست می دهد. . به طور خاص مجموعه داده GSE29105 را در اینجا بررسی میکنیم و توجه داریم که 4 مجموعه داده دیگر نیز مشابه همین مجموعه داده هستند.

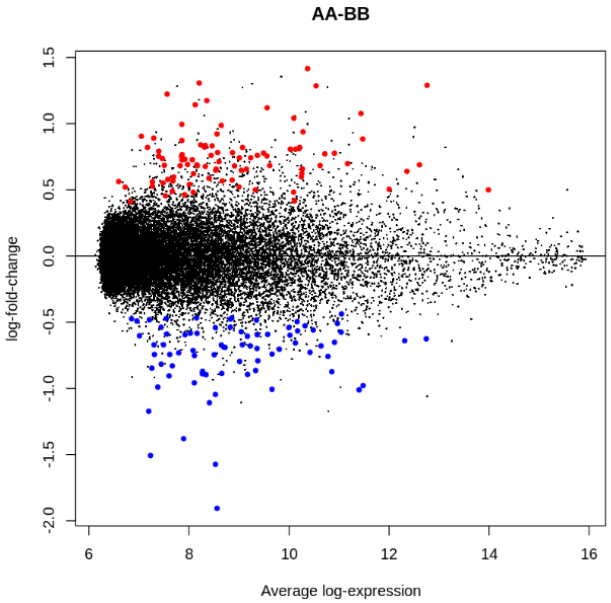
برای محاسبه نتایج ابتدا یک مدل خطی را روی مجموعه داده پیش پردازش شده fit میکنیم. سپس، با استفاده از مدل بیزی و آزمون های آماری، آماره ها و مقادیر لازم برای تحلیل را به دست می‌آوریم از قبیل P.Value، adj.P.Val، logFC و غیره. در نهایت نیز با استفاده از مقادیر به دست آمده و همچنین کشیدن نمودار های مختلف ژن هایی که در گروه های مختلف اهمیت دارند و differentially بین شده اند را شناسایی میکنیم. در ادامه در شکل های زیر نمودار توزیع مقادیر adj.P.Val، نمودار آتش فشانی و نمودار MD را مشاهده میکنید.



شکل 12.نمودار توزیع مقادیر adj.P.Val برای مجموعه داده GSE29105



شکل 13.نمودار آتش فشانی گروه AA در برابر BB برای مجموعه داده GSE29105



شکل 14.نمودار MD گروه AA در برابر BB برای مجموعه داده GSE29105

3-1-3- داده‌های افزوده مشابه مقاله اصلی به شکل ترکیب شده

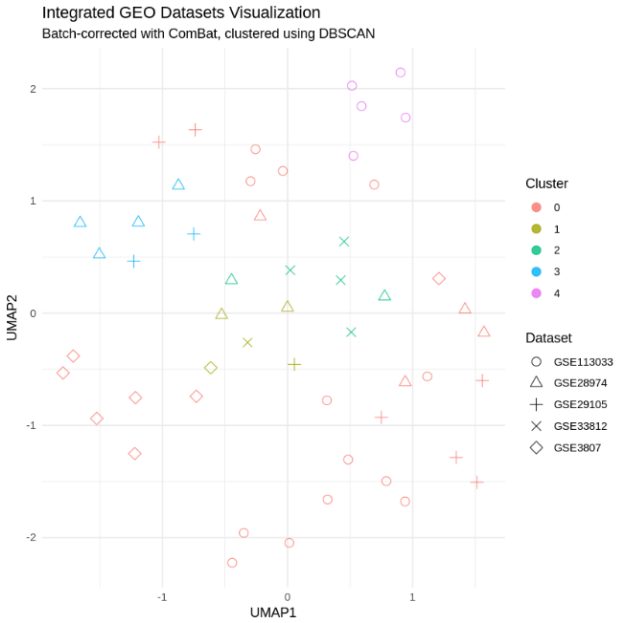
نمودار و خروجی های به دست آمده به همراه کد های لازم بـه شـکل کامـــل در فایـــل bio_proj_merged_5_datasets.ipynb قابل مشاهده است. در اینجا به توضیح مختصری از آن می پردازیم.

برای ارزیابی شباهت و اطمینان از کیفیت مجموعه داده های پیدا شـده با مجموعه داده اصلی موجود در مقاله و meta-analysis این بخش برای ما اهمیت دارد.

ابتدا از scale مناسـب داده هـای اطمینـان حاصـل میکنـیم چـرا کـه الگوریتم های بعدی از قبیل کاهش ابعاد به داده های نرمال شـده نیـاز دارند. سپس ژن های با واریانس صفر را حذف میکنـیم چـرا کـه هـیچ اطلاعات مفیدی برای ما ندارند.

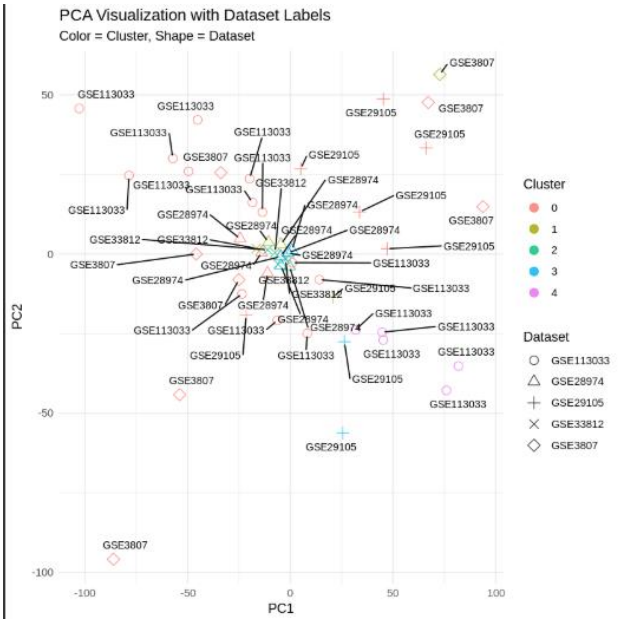
حال از الگوریم کاهش ابعاد UMAP با ابرپارامتر های مناسب استفاده میکنیم. سپس برای مصورسازی بهتر از یک الگوریتم خوشه بندی مانند DBSCAN استفاده میکنیم.

نتیجه به دست آمده را در شکل زیر مشـاهد میکنیـد. همـان طـور کـه میبینید نمونه های موجود در هر مجموعه داده به هم نزدیک تر هستند. همچنین مجموعه داده اصلی و مرجع ما یعنی GSE33812 در مرکز قرار گرفته که نشان دهنده نزدیکی نمونه های آن بـه سـایر مجموعـه داده ها است که گواه بر کیفیت و نزدیکی سایر مجموعه داده ها بـه آن هست. همچنین می توانیم مجموعه داده هـای همسـایه را ببینیـد کـه نشان می دهد کدام مجموعه داده ها قرابت بیشتری به هم دیگر دارند. نتیجه جالب و مورد انتظار دیگر اینکه مجموعـه داده GSE113033 از بقیـه فاصـله بیشـتری دارد کـه بـا توجـه بـه ماهیـت زیسـتی آن و توضیحات مجموعه داده کاملا مورد انتظار است.

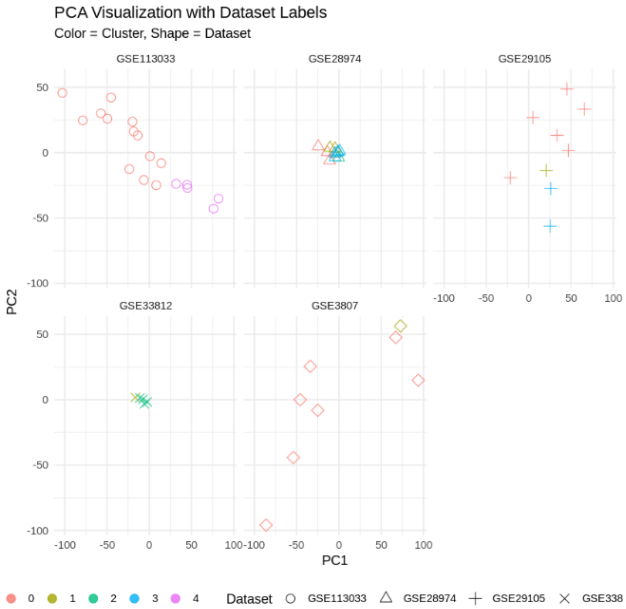


شکل 15.نمودار کاهش ابعاد با الگوریتم UMAP و اجرای الگوریتم خوشه بندی رو مجموعه داده های ترکیب شده با هم

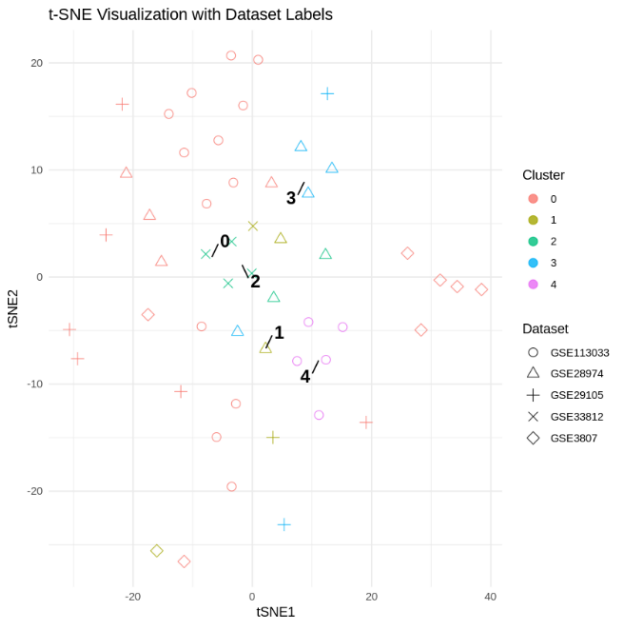
همچنین برای تحلیل جامع تر از الگوریتم های کاهش ابعاد PCA و t-SNE نیز استفاده کردیم و نمودار های زیر را کشیدیم که تایید کننده توضیحات و نتایج بالا هستند.



شکل 16.: نمودار کاهش ابعاد با الگوریتم PCA و اجرای الگوریتم خوشه بندی رو مجموعه داده های ترکیب شده با هم



شکل 17.نمودار کاهش ابعاد با الگوریتم PCA و اجرای الگوریتم خوشه بندی رو مجموعه داده های ترکیب شده با هم به تفکیک مجموعه داده



شکل 18.نمودار کاهش ابعاد با الگوریتم t-SNE و اجرای الگوریتم خوشه بندی رو مجموعه داده های ترکیب شده با هم

4- نتیجه گیری

دراین مطالعه، تغییرات بیان ژن در سلول‌های بنیادی مزانشیمی بیماران مبتلا به کم‌خونی آپلاستیک و افراد سالم مورد بررسی قرار گرفت. ابتدا روش انتخاب ژن‌های مورد مقایسه در این دو گروه تحلیل شد. سپس مسیرهای زیستی تحت تأثیر این ژن‌ها، فرآیندهای زیستی مرتبط و ویژگی‌های فنوتیپی بررسی شدند. نتایج نشان می‌دهد که سلول‌های بنیادی مزانشیمی نقش مهمی در پیشرفت کم‌خونی آپلاستیک دارند و تغییرات ژنتیکی در آن‌ها قابل شناسایی است. بیان ژن در این سلول‌ها بر مسیرهای مختلفی از جمله چرخه سلولی، تقسیم سلولی، تکثیر، مهاجرت سلولی، تمایز سلول‌های خونی تأثیر می‌گذارد که نشان‌دهنده‌ی این است نقص عملکرد سلولی به عنوان یکی از ویژگی‌های اصلی این بیماری است .

همچنین در مرحله دوم به علت تعداد کم نمونه ها در داده اصلی،5 مجموعه داده مشابه مجموعه داده اصلی مقاله را برای تحلیل بیشتر جمع آوری کردیم. سپس در مرحله بعد هر کدام را به طور مستقل از هم تحلیل کردیم و ژن های پر اهمیت آنها را شناسایی کردیم. در قدم بعد، این مجموعه داده ها را با مجموعه داده اصلی ترکیب کردیم و ژن های مشترک را کاهش ابعاد داده و الگوریتم خوشه بندی روی آنها اجرا کردیم. در نتیجه دیدیم که این مجموعه داده ها شباهت خوبی با مجموعه داده اصلی مقاله و کیفیت مناسبی دارند.

مراجع

[1] S.A. Peslak, T. Timothy Olson, D.V. Babushok, Diagnosis and treatment of aplastic anemia, Curr. Treat. Options Oncol. 18 (12) (2017) 70.

[2] N.S. Young, D.W. Kaufman, The epidemiology of acquired aplastic anemia, Haematologica 93 (4) (2008) 489–492.

[3] C. Dufour, P. Veys, E. Carraro, N. Bhatnagar, M. Pillon, R. Wynn, et al., Similar outcome of upfront-unrelated and matched sibling stem cell transplantation in idiopathic paediatric aplastic anaemia, Br. J. Haematol. 171 (4) (2015) 585–594.

[4] Kaufman DW, Kelly JP, Levy M, Shapiro S. The Drug Etiology of Agranulocytosis and Aplastic Anemia 1991; Oxford University Press, New York.

[5] L.E. Böttiger, B. Westerholm, Aplastic anaemia: I. Incidence and aetiology, Acta Med. Scand. 192 (1972) 315–318.

[6] S.M. Davies, D.J. Walker, Aplastic anaemia in the Northern Region 1971–1978 and follow-up of long term survivors, Clin. Lab. Haematol. 8 (1986) 307–313.

[7] M. Szklo, L. Sensenbrenner, J. Markowitz, S. Weida, S. Warm, M. Linet, Incidence of aplastic anemia in metropolitan Baltimore: a population-based study, Blood 66 (1985) 115–119.

[8] C. Isaragrisl, C. Sriratanasatavorn, A. Piankijagum, S. Vannasaeng, Y. Porapaklham, P.E. Leaverton, et al., Incidence of aplastic anemia in Bangkok, Blood 77 (1991) 2166–2168.

[9] N.S. Young, Aplastic anemia, Lancet 346 (1995) 228–232.

[10] N.S. Young, R. Calado, P. Scheinberg, Current concepts in the pathophysiology and treatment of aplastic anemia, Blood 108 (2006) 2509–2519.

[11] R.A. Brodski, R.J. Jones, Aplastic anaemia, Lancet 365 (2005) 1647–1656.

[12] N. Yoshida, H. Yagasald, A. Hama, Y. Takahashi, Y. Kosaka, R. Kobayashi, et al., Predicting response to immunosuppressive therapy in childhood aplastic anemia, Haematologica 96 (2011) 771–774.

[13] F. Timeus, N. Crescenzio, A. Lorenzati, et al., Paroxysmal nocturnal hemoglobinuria clones in children with acquired aplastic anemia: a prospective single centre study, Br. J. Haematol. 150 (2010) 483–485.

[14] L.C. Fox, D.S. Ritchie, Pediatric aplastic anemia treatment patterns and responses; power in the numbers, Haematologica 104 (10) (2019) 1909–1912.

[15] E.H. Blackburn, Switching and signaling at the telomere, Cell 106 (6) (2001) 661–673.

[16] H. Sakaguchi, N. Nishio, A. Hama, N. Kawashima, X. Wang, A. Narita, S. Doisaki, Y. Xu, H. Muramatsu, N. Yoshida, Y. Takahashi, K. Kudo, H. Moritake, K. Nakamura, R. Kobayashi, E. Ito, H. Yabe, S. Ohga, A. Ohara, S. Kojima, Japan Childhood Aplastic Anemia Study Group, Peripheral blood lymphocyte telomere length as a predictor of response to immunosuppressive therapy in childhood aplastic anemia, Haematologica 99 (8) (2014) 1312–1316.

[17] B. Hubner, S. Merk, S. Rauhut, M. Dugas, T. Haferlach, M. Fuehrer, A. Borkhardt, Individual gene expression profiling of bone marrow CD34 cells in acquired severe aplastic anemia (aSAA) in children, Blood 108 (11) (2006) 978.

[18] J. Li, S. Yang, S. Lu, H. Zhao, J. Feng, W. Li, F. Ma, Q. Ren, B. Liu, L. Zhang, et al., Differential gene expression profile associated with the abnormality of bone marrow mesenchymal stem cells in aplastic anemia, PLoS One 7 (2012) e47764.

[19] M.C. Kastrinaki, K. Pavlaki, A.K. Batsali, E. Kouvidi, I. Mavroudi, C. Pontikoglou, H.A. Papadaki, Mesenchymal stem cells in immune-mediated bone marrow failure syndromes, Clin. Dev. Immunol. 2013 (2013) 265608.

[20] Y.H. Chao, C.T. Peng, H.J. Ham, C.K. Chan, K.H. Wu, Poor potential of proliferation and differentiation in bone marrow mesenchymal stem cells derived from children with severe aplastic anemia, Ann. Hematol. 89 (2010) 715–723.

[21] E. Hamzic, K. Whiting, E. Gordon Smith, R. Pettengell, Characterization of bone marrow mesenchymal stromal cells in aplastic anaemia, Br. J. Haematol. 169 (2015) 804–813.

[22] S. Fujimaki, H. Harigae, T. Sugawara, et al., Decreased expression of transcription factor GATA-2 in haematopoietic stem cells in patients with aplastic anaemia, Br. J. Haematol. 113 (2001) 52–57.

[23] Y.H. Chao, K.H. Wu, S.H. Chiou, et al., Downregulated CXCL12 expression in mesenchymal stem cells associated with severe aplastic anemia in children, Ann. Hematol. 94 (2015) 13–22.

[24] Y. Xu, Y. Takahashi, Y. Wang, et al., Downregulation of GATA-2 and overexpression of adipogenic gene PPAR-

[39] S. Adbikari, P. Mandal, Integrated analysis of global gene and microRNA expression profiling associated with aplastic anaemia, Life Sci. 228 (2019) 47–52.
[40] "Reactome," http://www.reactome.org
gamma in mesenchymal stem cells from patients with aplastic anemia, Exp. Hematol. 37 (2009) 1393–1399.
[25] E.P. Weinzierl, D.A. Arber, The differential diagnosis and bone marrow evaluation of new-onset pancytopenia, Am. J. Clin. Pathol. 139 (1) (2013) 9–29.
[26] R.M. Cawthon, Telomere measurement by quantitative PCR, Nucleic Acids Res. 30 (10) (2002) e47.
[27] N. O'Callaghan, V. Dhillon, P. Thomas, M. Fenech, A quantitative real-time PCR method for absolute telomere length, Biotechniques 44 (6) (2008) 807–809.
[28] B. Sebastian, Protocol GMBOO3:B: 20090216NP, Genomic Medicine Biorepository GMB003, Revised and Approved 2012JULY26 by B. Sebastian.
[29] K. Chow, et al., Gene Expression Profiles of Bone Marrow Mesenchymal Stem Cells in Pediatric Patients with Severe Aplastic Anemia, 2011 (Chow K. et al., 2011, accession GSE33812).
[30] F. Beier, M. Foronda, P. Martinez, M.A. Blasco, Conditional TRF1 knockout in the hematopoietic compartment leads to bone marrow failure and recapitulates clinical features of dyskeratosis congenita, Blood 120 (15) (2012) 2990–3000.
[31] J.L. Cheng, A.L. Wang, J. Wan, Association between the M235T polymorphism of the AGT gene and cytokines in patients with hypertension, Exp. Ther. Med. 3 (3) (2011) 509–512.
[32] K. Matsushita, Y. Wu, Y. Okamoto, R.E. Pratt, V.J. Dzau, Local renin angiotensin expression regulates human mesenchymal stem cell differentiation to adipocytes, Hypertension (Dallas, Tex.: 1979) 48 (6) (2006) 1095–1102.
[33] W.X. Carroll, N.S. Kalupahana, S.L. Booker, N. Siriwardhana, M. Lemieux, A.M. Saxton, et al., Angiotensinogen gene silencing reduces markers of lipid accumulation and inflammation in cultured adipocytes, Front. Endocrinol. 4 (2013) 10.
[34] S. Stopp, M. Grindl, M. Fackler, J. Malkmus, M. Leone, R. Naumann, S. Frantz, E. Wolf, B. von Eyss, F.B. Engel, S. Gaubatz, Deletion of Gas2l3 in mice leads to specific defects in cardiomyocyte cytokinesis during development, Proc. Natl. Acad. Sci. U.S.A. 114 (2017) 8029–8034.
[35] S. Kordasti, B. Costantini, T. Seidl, P.P. Abellan, M.M. Llordella, D. McLornan, K. Diggins, A. Kulasekararaj, C. Benfatto, X. Feng, A. Smith, S.A. Mian, R. Melchiotti, E. de Rinaldis, R. Ellis, N. Petrov, G.A.M. Powles, S.S. Chung, N.S. Thomas, F. Farzaneh, J.M. Irish, S. Heck, N.S. Young, J.C.W. Marsh, G.J. Mufti, Deep phenotyping of Tregs identifies an immune signature for idiopathic aplastic anemia and predicts response to treatment, Blood 128 (9) (2016) 1193–1206.
[36] P. Bertheau, E. Turpin, D.S. Rickman, M. Espie, A. de Reynies, J.P. Feugeas, L. Plassa, H. Soliman, M. Varna, A. de Roquancourt, J. Lehmann-Che, Y. Beuzard, M. Marty, J.L. Misset, A. Janin, H. de The, Exquisite sensitivity of TP53 mutant and basal breast cancers to a dose-dense epirubicin-cyclophosphamide regimen, PLoS Med. 4 (2007) e90.
[37] C.N.N. Weiss, K. Ito, A macro view of MicroRNAs: the discovery of MicroRNAs and their role in hematopoiesis and hematologic disease, Int. Rev. Cell Mol. Biol. 334 (2017) 99–175.
[38] E.C. Guinan, Aplastic anemia: management of pediatric patients, Hematology Am. Soc. Hematol. Educ. Program (2005) 104–109.