آنالیز داده های ژل های الکتروفورز دو بعدی با استفاده از روش های آماری چند متغیره

حکیمه زالی[۱]، مصطفی رضایی طاویرانی[۲*]، علی صیدخانی نهال[۳]، مهدی شهریاری نور[٤]، نازنین بلغاری[٤]

۱)دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی

۲)مرکز تحقیقات پروتئومیکس، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی تهران

۳)گروه بیوشیمی بالینی، دانشکده پزشکی، دانشگاه علوم پزشکی ایلام

۴) گروه میکروبیولوژی، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات گیلان

تاریخ دریافت: ۹۱/۱۰/۱۲ **تاریخ پذیرش: ۹۱/۱۲/۹**

چکیده

مقدمه : آنالیز داده های عظیم پروتئومیکی با دارا بودن متغیر های زیاد نیاز به روش های چند متغیره است که امکان آنالیز آماری همزمان چندین متغیر را فراهم می کنند. در این مطالعه فرآیند تمایز و پیری سلولی از سلول های بنیادی تا آستروسیت توسط پروتئومیکس مورد بررسی قرار گرفته و کاربردهای آماری نرم افزار های آنالیز ژل مورد توجه قرار گرفت.

مواد و روش ها: پروتئوم چهار گروه آزمایشی (سلول های بنیادی، آستروسیت جوان ، آستروسیت با تمایز متوسط و آستروسیت پیر) توسط نرم افزار Progenesis Same spots مورد مطالعه آماری قرار گرفت . آنالیز خوشه بندی ،آنالیز مولفه اصلی و آنالیز Power در بررسی گروه های آزمایشی مورد استفاده قرار گرفت.

یافته های پژوهش: در آنالیز بیوانفورتیکی و آماری ژل های بدست آمده از تکنیک الکتروفورز دو بعدی ۹۴۰ نقطه پروتئینی با تغییر بیان معنی دار ($p<0.05$)در چهار گروه مشاهده شد که مقایسه بین گروه ها حکایت از بیان پروتئینهای جدید و خاموشی برخی از پروتئین ها در مسیر های سیگنالینگ تمایز و پیری سلولی است.

بحث و نتیجه گیری: آنالیز خوشه بندی، پروتئین ها را از نظر بیان به ۲ خوشه اصلی تقسیم نمود که بیانگر وجود پروتئین هایی با بیان مشابه در هر خوشه است که این پروتئین ها می توانند عملکرد مشابهی را در شرایط آزمایش ارائه نمایند یا بیانگر حضور همه آن ها در مسیر بیولوژیکی مشترکی است . آنالیز PCA نتایج حاصل از خوشه بندی را تایید نمود و نشان داد که داده های پروتئینی بر طبق شرایط آزمایش خوشه‌بندی شده است. در نهایت می توان نتیجه گرفت که با کمک نرم افزار ،آنالیز آماری به سادگی و با سرعت انجام شد و تغییرات بیان معنی دار ناشی از تمایز و پیری در سطح پروتئوم با آنالیزهای آماری خوشه بندی و PCA به خوبی مورد بررسی قرار گرفت و شاخص های تغییرات مشخص شد.

واژه های کلیدی: ژل الکتروفورز دو بعدی ، روش آماری چند متغیره ، پروتئین

***نویسنده مسئول:** مرکز تحقیقات پروتئومیکس، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی تهران

Email: rezaei.tavirani@ibb.ut.ac.ir

[Downloaded from sjimu.medilam.ac.ir on 2026-11-11]

مقدمه

الکتروفورز دو بعدی(Two Dimensional Electrophoresis, 2DE) یکی از رایج ترین و قدرتمندترین روشهای جداسازی پروتئینها در پروتئومیکس است. این روش در حدود سی سال پیش با بهره گیری از تفاوت موجود در دو خصوصیت ذاتی پروتئین ها, یعنی بار الکتریکی و وزن مولکولی, ابداع شد(۱–۳).در سال های اخیر تکنولوژی مربوط به توصیف لکه های پروتئینی نمایان شده بر ژل های الکتروفورز دو بعدی، توسعه ی قابل ملاحظه ای یافته و بانک های اطلاعاتی متعددی نیز در این زمینه ایجاد و گسترش یافته است که تأثیر این ابتکارات بر ارتقاء پروتئومیک بسیار چشمگیر است. تشخیص زودرس بیماری ها، یکی از آروزهای دیرینه ی بخش بالینی است که فناوری پروتئومیکی با فراهم آوردن این امکان، نوید بخش درمان سریع و مؤثر بیماری های مهلک می باشد.استفاده از بیوانفورماتیک و بانک های اطلاعاتی به منظور تفسیر یافته های پروتئومیکی و نیز فراهم آوردن امکان استفاده از دستاوردهای این فناوری به طور شگرفی در حال توسعه می باشد. (۴–۸) .آنالیز آماری داده های عظیم پروتئومیکی با دارا بودن متغییر های زیاد نیاز به روش های چند متغیره است که امکان آنالیز آماری همزمان چندین متغیر را فراهم می کنند . مسائل مربوط به یک، دو و یا سه متغیر را می توان بطور ذهنی تصور کرد و یا بطور گرافیکی نمایش داد ولی گاهی در مسائل اکتشافی با یک فضای ۱۰، ۲۰ و یا ۵۰ متغیره روبرو هستیم که بررسی روابط بین آنها را بسیار دشوار می سازد. در اینگونه موارد لازم است با استفاده از روش های آماری چند متغیره به کاهش تعداد بعدها در فضای مورد بررسی پرداخت، بطوریکه نتایج این ابعاد جدید (متغیرهای جدید) با تعدادی به مراتب کمتر از حالت قبل، بتواند بخش اعظم تغییر پذیری داده ها را تشریح کنند. نکته ای که در آمار چند متغیره باید به آن توجه شود، تعداد نمونه ها در جوامع تحت بررسی است. معمولاً روش های چند متغیره نیازمند تعداد زیادی نمونه است. اعتبار این تحلیل ها تا حدودی تابع بزرگی جامعه نمونه تحت بررسی است .امروزه نرم افزار های آنالیز ژل همه آنالیزهای آماری مورد نیاز در آنالیز آماری برای ژل های دو بعدی را انجام می دهند و در واقع این نرم افزار ها به طور عمیقی به دنبال دستیابی به هدف های ویژه ای هستند از جمله کشف بیومارکرها با تعیین میزان بیان پروتئین های منفرد، جداسازی یک یا تعداد بیشتری نقاط پروتئینی در تصویر ژل اسکن شده و مطابقت دادن نقاط در ژل ها در نمونه های مشابه جهت تعیین تفاوت پروتئومیکی بین مراحل اولیه و پیشرفته یک بیماری. نرم افزارهای آنالیز ژل های دو بعدی از دو روش آماری اصلی ،طراحی بین موضوعی (Between – subject design) و طراحی درون موضوعی (Within subject design)بهره می جویند.آنالیز اجزای اصلی، خوشه بندی، آنالیز همبستگی و آنالیز Power روش های آماری چند منغیری هستند که در آنالیز ژل های دو بعدی استفاده می شود. اگر ما در جایی مجبور هستیم مهم ترین متغیر را یا یک تعداد محدودی از متغیرها را در یک مجموعه انتخاب کنیم از آنالیز اجزای اصلی کمک می گیریم. خوشه بندی به دسته بندی پروتئین ها بر اساس معیار مورد نظر می پردازد. آنالیز همبستگی روابط بین پروفایل بیانی نقاط پروتئینی را ارزیابی می کند. با توجه به اینکه داده های پروتئومیکی دارای متغیرهای متعددی است لذا استفاده از روش آنالیز آماری چند متغیره که می تواند همزمان همه نقاط پروتئینی موجود را مورد بررسی قرار دهد و تفاوت های کلی بیان ژن را بین گروه های مختلف درگیر در آنالیز را شناسایی نماید روش موثری است(۹–۱۳). در این مطالعه زل های دو بعدی حاصل از تمایز استم سل به سلول عصبی و مراحل مختلف تبدیل یک عصب جوان به پیر مورد بررسی قرار می گیرد . در کل ۴ گروه مطالعاتی(استم سل ،سلول آستروسیت جوان(تمایز کم)، سلول آستروسیت با تمایز متوسط و سلول آستروسیت پیر وجود دارد. تعداد نقاط پروتئینی و تفاوت بیان بین آن ها توسط نرم افزار nonlinear progenesis same spot مورد آنالیز قرار می گیرد.

مواد و روش ها

[Downloaded from sjimu.medilam.ac.ir on 2026-11-11]

نمونه گیری، استخراج پروتئین و الکتروفورز دو بعدی:

چهار گروه آزمایشی شامل رده سلولی تمایز نیافته (سلول های بنیادی)وسلول های تمایز یافته (سلول های آستروسیت جوان) و سلول های آستروسیتی که به فاصله ۲ هفته (آستروسیت با تمایز متوسط) و ۴ هفته (آستروسیت پیر) برداشت شدند. از همه گروه ها پروتئین استخراج شد.غلظت پروتئینهای چهار گروه سلولی به روش برادفورد مورد سنجش قرار گرفت و برای هر نمونه سه ژل های الکتروفورز دو بعدی تکرار صورت پذیرفت و در نهایت رنگ آمیزی و آنالیز ژل انجام شد.

آنالیز تصویر ژل‌های دوبعدی:

الکتروفورز دوبعدی قادر است صدها پلی پپتید را از یک محلول پروتیئنی بر اساس بار الکتریکی و وزن مولکولی آنها جدا کند. پس از طی مراحل رنگ‌آمیزی، پروتئین‌ها به صورت لکه‌هایی با صفات و ویژگی‌های ظاهری متفاوت مانند شکل، اندازه و شدت رنگ نمایان می‌گردند. پس از اسکن ژل ، استخراج داده‌های مربوط به نمونه بافت هیپوکامپ موش های دو گروه با نرم افزار Non Linear Progenesis Same Spot مورد بررسی قرار گرفت. تعداد نقاط پروتئینی و تفاوت آن ها (افزایش و یا کاهش بیان) توسط این نرم افزار آنالیز شد و با آمار چند متغیره(آنالیزمولفه اصلی) پروتئین ها از نظر بیان خوشه بندی شدند (۱۲-۱۵).

روش‌های تجزیه و تحلیل داده‌ها:

طراحی آزمایش برای تصاویر ژل ها شامل دو روش طراحی است ۱- طراحی بین موضوعی (Between – subject design) که نمونه‌هایی از هر موضوع تحت یک شرایط آورده شده است (برای مثال کنترل بر علیه نمونه‌های تحت داروهای متفاوت). محاسبهٔ ANOVA بدین ترتیب است که شرایط مستقل هستند و تست آماری بر این فرض است که میانگین این شرایط نیز مساوی است.

۲- طراحی درون موضوعی (Within subject design) ، نمونه‌ها از یک موضوع تحت شرایط متفاوت (برای مثال موضوع مشابه در دوره‌های زمانی یا بعد از یک یا تعداد بیشتری تیمار) گرفته می‌شود. در اینجا یک ANOVA استاندارد به عنوان داده‌های فرض ANOVA مستقل کافی به نظر نمی‌رسد بنابراین با استفاده از repeated measures ANOVA تفاوتهای فردی می‌تواند حذف شود یا به عنوان یک منبع بین شرایط مختلف کاهش یابد. در طراحی درون موضوعی می‌توان به یک paired samples t-test گسترده که شامل مقایسهٔ بین تعداد بیشتری از دو repeated measures است اندیشید. با توجه به ماهیت گروه بندی نمونه های مورد بررسی ،در اینجا از طراحی بین موضوعی برای آنالیز های مقایسه ای استفاده می شود.

با استفاده از آنالیز همبستگی، ارزیابی روابط بین پروفایل بیانی نقاط انتخاب شده صورت می‌گیرد. سپس این پاسخ به صورت گرافیکی به شکل دندوگرام فعال که فاصله‌های عمودی بین جفت نقاط بیانگر میزان شباهت پروفایل بیانی آنهاست. آنالیز همبستگی قادر به گروه بندی نقاط با همدیگر بر طبق میزان شباهت پروفایل بیانی آنهاست. بنابراین می‌تواند نشان دهد که کدام یک ممکن است با هم تنظیم شوند و به عبارت دیگر شاید در مسیری بیولوژیکی مشابه در کنار هم قرار گرفته باشند.

PCA می‌تواند جهت تعیین اینکه هیچ Outlier در داده‌ها وجود ندارد و همچنین می‌تواند مشخص نماید که چقدر نمونه‌ها خوب دسته بندی شده‌اند. گروه‌بندی را می‌توان در نمودارهای PCA دو بعدی جهت مقایسه‌ی گروه‌بندی طراحی آزمایش مشاهده نمود. محاسبات بر طبق outlierهای ممکن در گروه-بندی اولیه انجام می‌پذیرد.

Power یک تست آماری انعکاسی از در توانایی داده‌های آزمایش شده جهت یافتن تفاوتهای است که واقعاً وجود دارند.Power به صورت درصد بیان می‌شود و Power ۸۰٪ یعنی یک سطح پذیرفته شده که به شما اجازه می‌دهد که تعداد تکرارهای نمونه که باعث حصول Power ۸۰٪ می‌شود، ارزیابی نمایید.

یافته های پژوهش

نرم افزار های آنالیز ژل با استفاده از آمار به دنبال جداسازی یک یا تعداد بیشتری نقاط پروتئینی در تصویر ژل اسکن شده و مطابقت دادن نقاط در ژل ها

[Downloaded from sjimu.medilam.ac.ir on 2026-11-11]

در نمونه های مشابه جهت تعیین تفاوت پروتئومیکی بین حالت های سالم و بیماری یا مراحل مختلف یک بیماری و کشف بیومارکرها با تعیین میزان بیان پروتئینی آن ها هستند. در جدول ۱ نرم افزار هایی که امروزه جهت آنالیز ژل های دو بعدی استفاده می شود نشان داده شده است. از نرم افزارهای آنالیز ژل نرم افزار Progenesis Same spots انتخاب شد و ژل های الکتروفورز دو بعدی چهار گروه شرکت کننده در آزمایش توسط آن مورد آنالیز قرار گرفت.

شکل ۱ ژل های آستروسیت جوان (Condition۱) ، آستروسیت با تمایز متوسط(Condition۲)،آستروسیت پیر (Condition۳) و استم سل (Condition۴) را نمایش می دهد.

شکل ۲ صفحه طراحی آزمایش را نشان می دهد. نمونه های مورد بررسی در اینجا از طراحی بین موضوعی برای آنالیز های مقایسه ای بعدی استفاده شد.

نتایج حاصل از مقایسه ژل های چهار گروه توسط نرم افزار در شکل ۳ آمده است . از 1469 نقطه پروتئینی ۹۴۰ نقطه دارای تغییر بیان دو برابر داشتند که با انتخاب تصادفی تعدادی از نقاط جدا شده و درجدول ۲ آورده شده است. جدول میزان تغییر بیان (Fold)، میزان معنی داری بیان تغییر یافته(Anova (p))، pI وMW هر پروتئین و میزان بیان هر نقطه پروتئینی در گروههای آزمایشی را نشان می دهد.

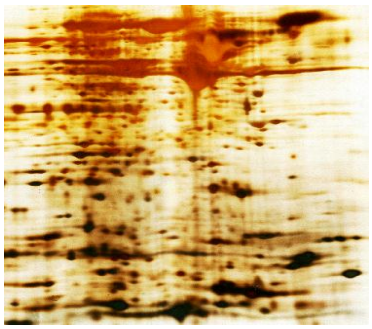
شکل ۴ آنالیز همبستگی نقاط پروتئینی را در ژل های مورد آزمایش نشان می دهد حاصل این آنالیز دو خوشه اصلی پروتئینی می باشد.

شکل ۵ آنالیز PCA دو بعدی را در نمونه های مورد آزمایش نشان می دهد.

شکل ۶ نمایش گرافیکی Power برای نمونه های مورد آزمایش را نشان می دهد. نمایش گرافیکی آن حکایت از ۷۲/۲٪ از نقاط استفاده شده در آزمایش است که Power ۸۰٪ یا بیشتر را می‌دهد و چگونگی ارتباط با تعداد تکرار را بیان می‌نماید. برای این Power حدود ۳۰ تکرار لازم است.

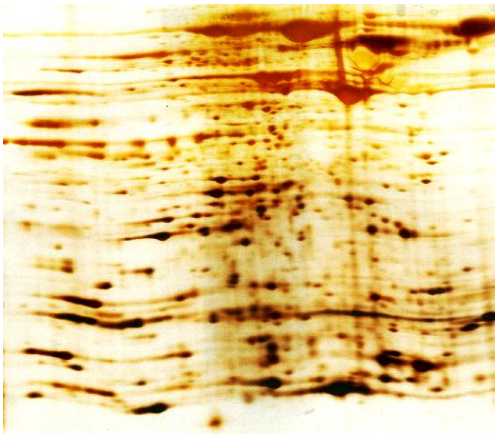
جدول ۱: نرم افزار هایی آنالیز ژل های دو بعدی مورد دسترس کاربران

Name	Formats Supported	Developers
BioNumerics 2D	2-D gel TIFF files	Applied Maths
Delta2D	2-D gel TIFF files	Decodon
ImageMaster 2D	2-D gel TIFF files	GE Healthcare
Melanie	2-D gel TIFF files	Geneva Bioinformatics (GeneBio)
PDQuest	2-D gel TIFF files	Bio-Rad Laboratories
Progenesis Samespots	2-D gel TIFF files	Nonlinear Dynamics
REDFIN	2-D gel TIFF files	Ludesi

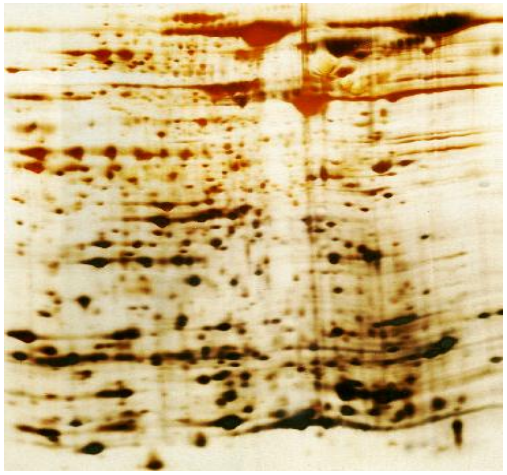


الف

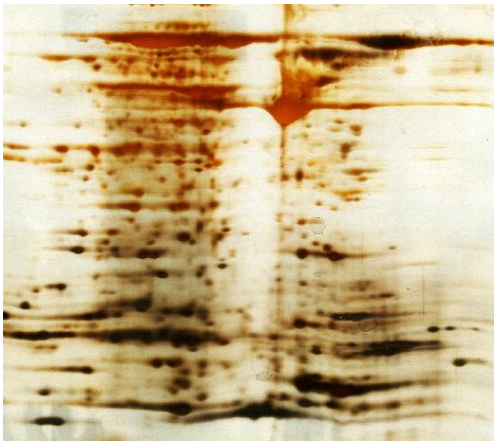
[Downloaded from sjimu.medilam.ac.ir on 2026-11-11]



ب



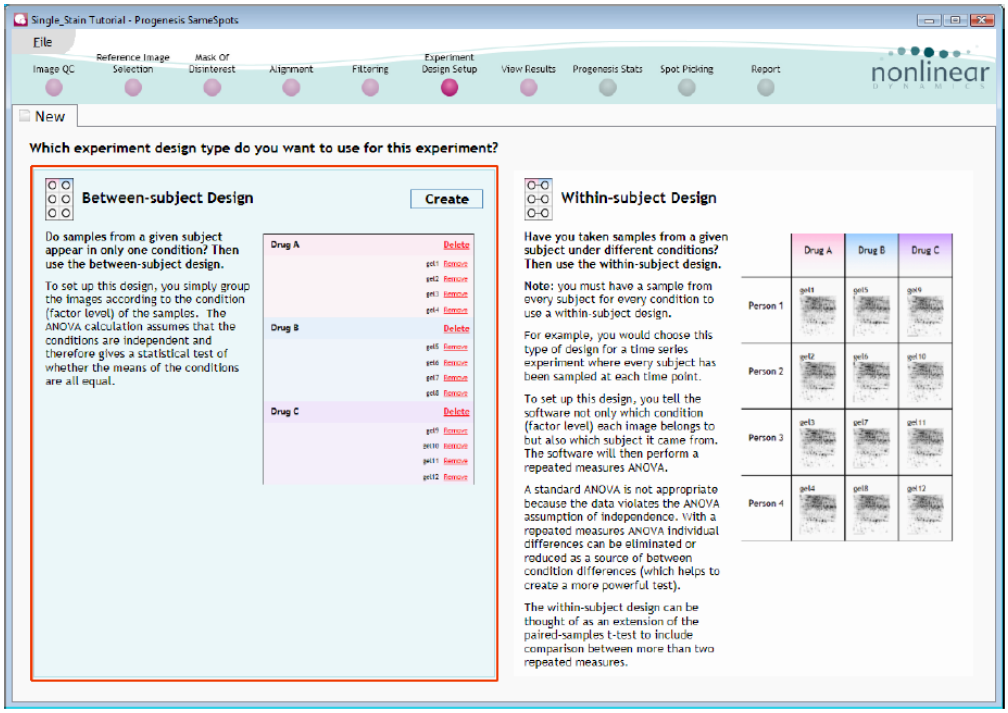
پ



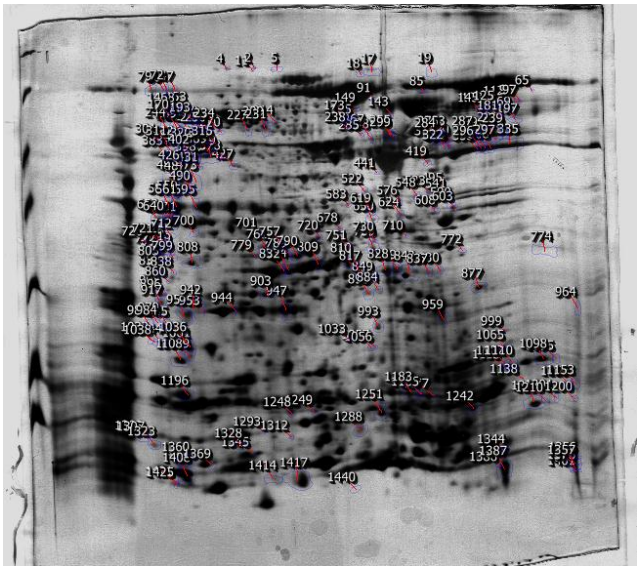
ج

شکل۱: شکل ظاهری ژل های دو بعدی استم سل (الف)،آستروسیت جوان(ب) ،
آستروسیت با تمایز متوسط (پ)و آستروسیت پیر (ج)

[Downloaded from sjimu.medilam.ac.ir on 2026-11-11]



شکل ۲: طراحی آزمایش بر اساس طراحی بین گروهی

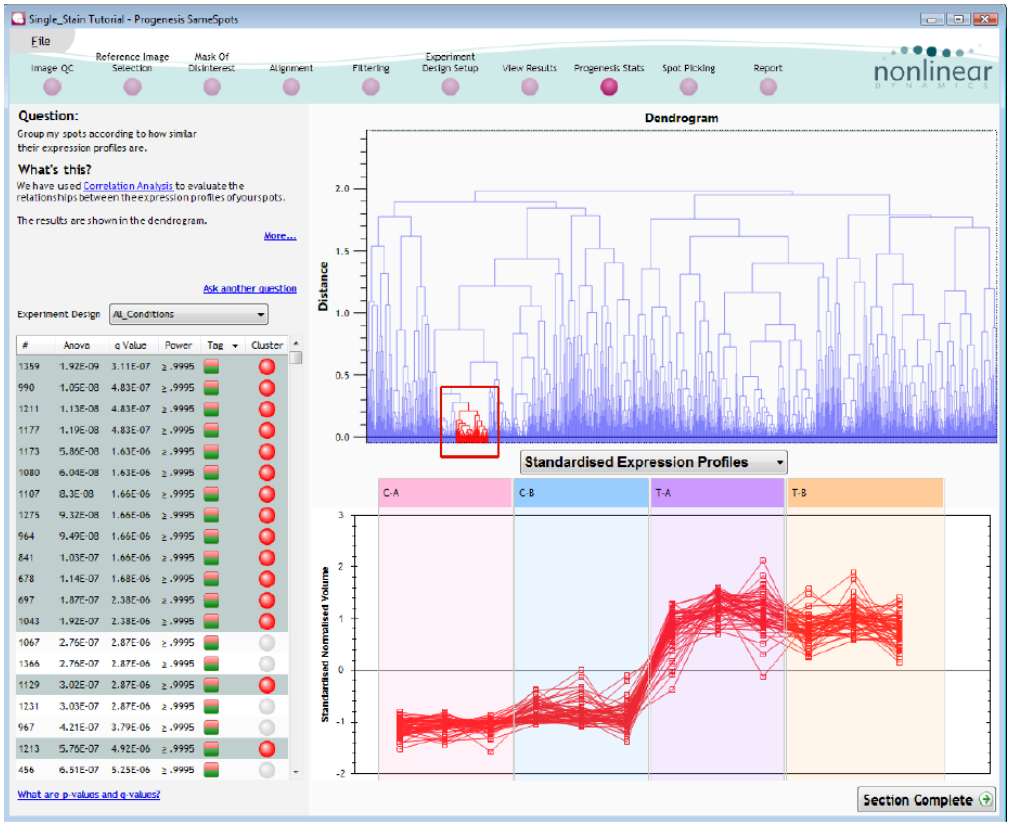


شکل ۳: نقاط مورد مقایسه در چهار گروه توسط نرم افزار progenesis
(اطلاعات تعدادی از نقاط در جدول ۲ آورده شده است)

[Downloaded from sjimu.medilam.ac.ir on 2026-11-11]

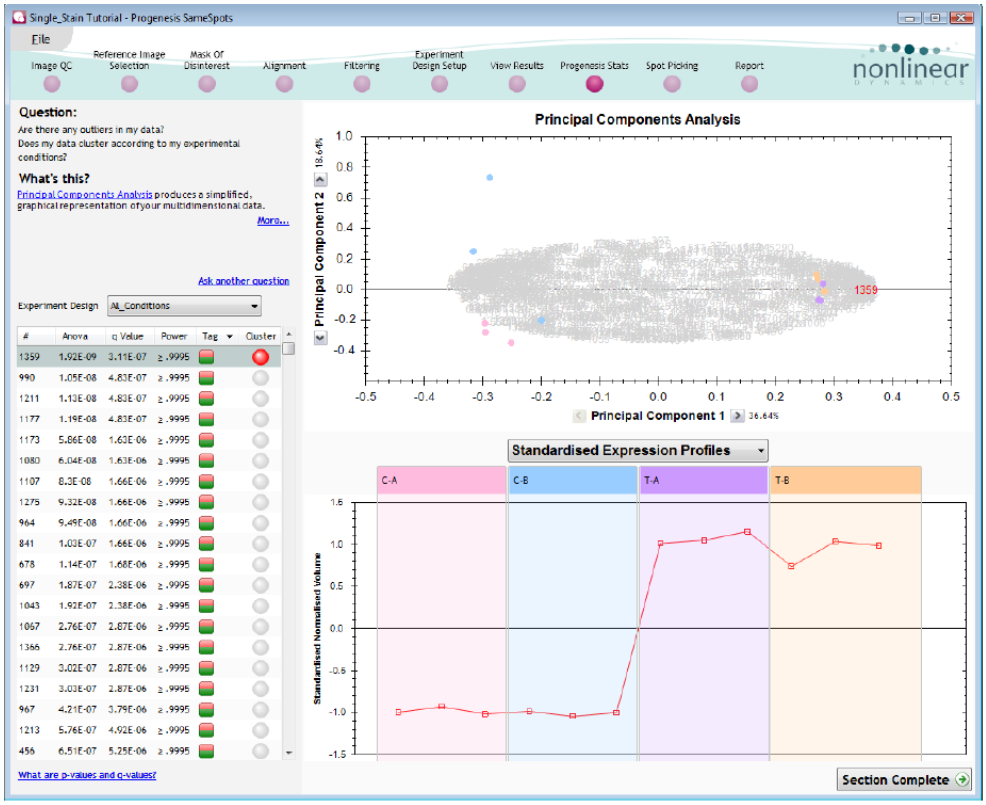
جدول ۲: آنالیز حاصل از مقایسه میزان بیان ژل های آستروسیت جوان (Condition۱) ، آستروسیت با تمایز توسط(Condition۲)،آستروسیت پیر (Condition۳) و استم سل (Condition٤) با نرم افزار progenesis (به علت حجم زیاد داده ها، مشخصات همه ژن های آنالیز شده در جدول آورده نشده). شماره هر نقطه پروتئینی(#)، میزان تغییر بیان (Fold)، میزان معنی داری بیان تغییر یافته((Anova (p)، pI وMW هر نقطه پروتئین و میزان بیان هر نقطه پروتئینی در گروههای آزمایشی (Average Normalised Volumes)

#	Anova (p)	Fold	pI	MW	Average Normalised Volumes			
					Condition 1	Condition 2	Condition 3	Condition 4
1249	0	4.2	6.54	25,362	5912.725	1.05E+04	2510	3155.24
661	0	2.1	4.74	64,409	6.15E+04	1.27E+05	6.78E+04	1.16E+05
1242	0	3.1	8.58	26,000	2.94E+04	2.69E+04	9480	1.48E+04
1344	0	3.1	9	18,771	9119.406	1.07E+04	3445	4986.186
300	0	3.1	8.28	104,295	1877.642	4264.783	1362	2812.667
1159	0	5	9.85	30,229	1.32E+04	3.94E+04	7831	8640.525
352	0	3.1	8.67	99,933	1.10E+04	3.46E+04	1.55E+04	3.09E+04
624	0	2.6	7.65	66,243	9025.474	1.77E+04	6731	1.11E+04

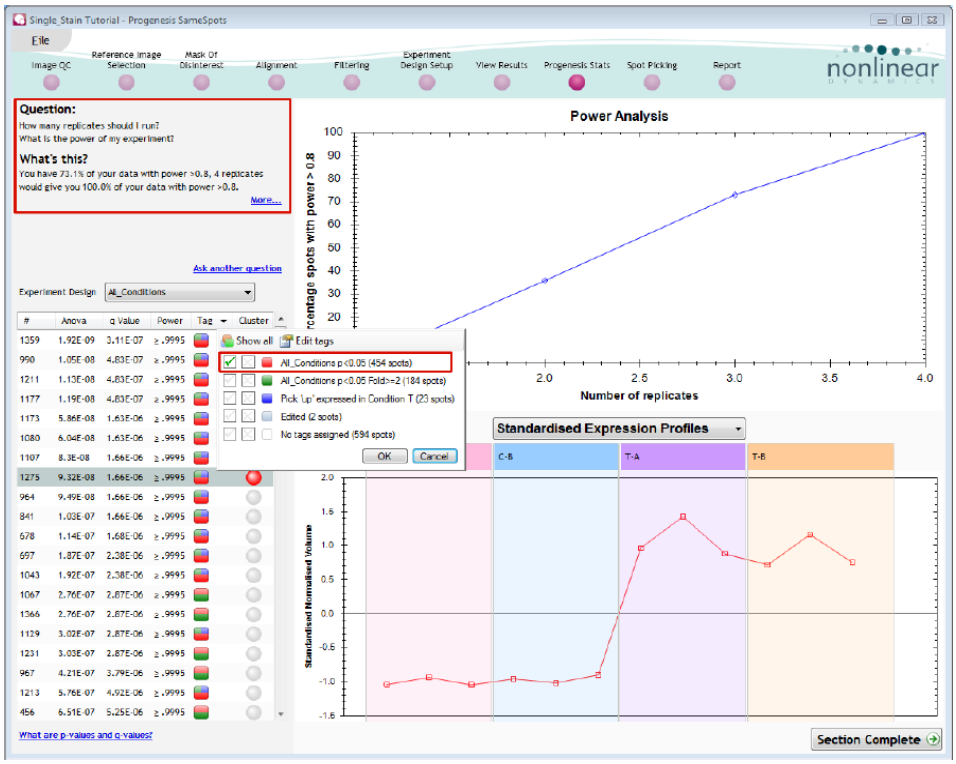


شکل ٤: آنالیز همبستگی نقاط پروتئینی در ژل های مورد آزمایش

[Downloaded from sjimu.medilam.ac.ir on 2026-11-11]



شکل ۵: آنالیز PCA را در نمونه ها مورد آزمایش



شکل ٦ نمایش گرافیکی Power برای نمونه های مورد آزمایش

[Downloaded from sjimu.medilam.ac.ir on 2026-11-11]

بحث و نتیجه گیری

نخستین نیاز در آنالیز پروتئوم جداسازی و آشکارسازی پروتئین های موجود در مخلوط هایی پیچیده (تا چند صد هزار پروتئین) می باشد. با اینکه امروزه روشهای جداسازی متعددی برای این منظور به کار گرفته شده است (از جمله روش های تراشه پروتئین (ProteinChip) و مطالعه مستقیم مخلوطهای پروتئینی به وسیله طیف سنجی جرمی و برچسب های تمایلی نشاندارشده با ایزوتوپ)، الکتروفورز دوبعدی(DE۲) روی ژل پلی اکریل آمید همچنان به عنوان یکی از بهترین گزینه ها برای مطالعه مخلوطهای پیچیده پروتئینی مطرح است. این امر به خاطر ویژگی استثنایی الکتروفورز دوبعدی در جداسازی همزمان چندین هزار پروتئین و به دنبال آن امکان آشکارسازی و مقایسه کمی پروتئین های جدا شده با حساسیت بالاست. در ضمن می توان با روشهای موجود مانند طیف سنجی جرمی پروتئین های جدا شده به وسیله ژل را شناسایی کرد(۲و۸). در این مطالعه روش های آماری مورد استفاده در آنالیز ژل های الکتروفورز دو بعدی مورد بررسی قرار گرفته شد. امروزه استفاده از نرم افزار های آنالیز ژل کمک بیوانفورماتیک را در ساده نمودن آنالیز داده های پروتئومیکی امری اجتناب ناپذیر نموده است. پیشرفت هایی که در سطح نرم افزاری صورت گرفته است آنالیز های آماری چند متغیره را در داده های چند بعدی پروتئومیکی تسهیل نموده است. جدول ا لیستی از نرم افزارهایی که امروزه در آنالیز ژل مورد استفاده قرار می گیرند آورده شده است. اساس آنالیز آماری همه این نرم افزارها جهت آنالیز ژل های الکتروفورز دو بعدی دو روش آماری اصلی است. روش های فرضیه محور که در ساده ترین حالت، آزمایش مقایسه دو نمونه های A و B است که برای یافتن آن دسته از پروتئین ها که نشان می دهد تفاوت های قابل توجهی در سطوح بیان دارد صورت می پذیرد. به طور معمول با چند تکرار در هر نمونه ، آزمون آماری Test Student's t برای پیدا کردن پروتئین های متفاوت بیان شده کاربرد دارد. از آزمون های ناپارامتریک آزمون های two-sample Wilcoxon test و Mann–Whitney U test در این مورد شناخته شده است. در آنالیز داده های ژل های الکتروفورز دو بعدی تعداد زیادی از آزمون ها می توانند تعداد بالای مثبت کاذب تولید کنند. به عنوان مثال، در یک آزمایش با ۲،۰۰۰ نقطه در ژل، نرخ پذیرفته آلفا مثبت کاذب ۵٪ در ۱۰۰ پروتئین یافت می شود . با وجودی که این تفاوت در نتیجه احتمال صرف است برای غلبه بر این مشکل تنظیم آلفا یک راه مناسب است. اصلاح Bonferroni کنترل احتمال خطا است . با این حال، این روش معمولا بیش از حد محافظه کارانه است .کنترل نرخ کشف کاذب(FDR) رویکردی است که می تواند موثر باشد نرخ کشف کاذب میزان نتایج مثبت کاذب در میان تمام پروفایلی است که مثبت بوده است. برای طرح های پیچیده تر آزمایش شامل عوامل چندگانه، ANOVA می تواند مورد استفاده قرار گیرد (۱۶–۲۰). روش دیگری که در جهت آنالیز های ژل های الکتروفورز دو بعدی به کار گرفته می شود روشهای مستقل از فرضیه(Hypothesis-independent methods) است که کشف الگو در مقادیر زیادی از داده ها ی چند بعدی است که در زمینه داده کاوی و یادگیری ماشین است.از آنجا که فرآیندهای بیولوژیکی در الگوهای بیانی از تعداد زیادی از پروتئین تشکیل شده اند روش های دوم برای تجزیه و تحلیل آزمایش های ژل های الکتروفورز دو بعدی موثر تر به نظر می رسد. خوشه بندی سلسله مراتبی اشاره به یک گروه از روش هایی دارد که هدف پروفایل های بیان گروهی است که تجزیه و تحلیل می تواند بیشتر روی آن ها انجام شود. خوشه بندی سلسله مراتبی از ژل را می توان برای شناسایی کردن outliers و برای شناسایی ساختار در آزمایش مورد استفاده قرار داد. خوشه بیانی از پروفیل ها برای شناسایی پروتئین هایی با رفتار مشابه انجام شده است که دلالت بر تنظیم با هم (co-regulat) آنها است(۱۹–۲۳).

تغییر ات بیان در سطح پروتئوم با کمک تکنیک ژل الکتروفورز دو بعدی مورد آنالیز بیوانفورماتیکی و آماری توسط نرم افزار progenesis same spot

[Downloaded from sjimu.medilam.ac.ir on 2026-11-11]

قرار گرفت . شکل ۱ ژل های چهار گروه آزمایشی را نشان می دهد . همانطور که در شکل ۲ نشان داده شده است طراحی آزمایش بر اساس روش بین گروهی بود .نتیجه آنالیز با نرم افزار در شکل ۳ مشخص است و بیان می دارد که ۹۴۰ نقطه پروتئینی در چهار گروه مشاهده شده است که مقایسه بین ژل ها در سطح بیان پروتینی حکایت از بیان نقاط پروتئینی جدید و خاموشی بیان برخی نقاط دیگر است. این نتیجه بیانگر این است که تمایز باعث تغییراتی در سطح بیان پروتئینی می شود. جدول ۲ برخی از نقاط پروتئینی در ژل را با مشخصات کامل به نمایش گذاشته است. نرم افزار ها در تعیین میزان بیان نقاط از میزان رنگی که لکه پروتئینی در آن جایگاه به خود گرفته است استفاده می نماید بنابراین یک خطا در نتایج حاصل از این نرم افزارها می تواند رخ دهد که باعث می شود کاربر را مجبور نماید که به صورت دستی نیز لکه ها را از نظر میزان بیان کنترل نماید و صرفا به نتایج حاصل از نرم افزار اعتماد ننماید. این خطا زمانی رخ می دهد که یک نقطه در یک ژل وجود نداشته باشد به عبارتی دیگر بیان نشده باشد اما نرم افزار به علت عدم رنگ بری یکسان ژل ها به اشتباه میزان بیانی را با توجه به زمینه ژل ارائه نماید. نرم افزار progenesis same spot نیز از این خطا مستثنی نیست و نیاز به کنترل همه نقاط در نتیجه نهایی به صورت دستی وجود دارد. اگر رنگ بری همه ژل ها یکسان باشد میزان این خطا به حداقل می رسد.

نرم افزار progenesis same spot با استفاده از روش آنالیز همبستگی به خوشه بندی پروتئین ها بر اساس میزان بیان آن ها می پردازد ، به طوری که پروتئین هایی که در یک خوشه دسته بندی می شوند، دارای بیان مشابهی می باشند ؛ بدین معنی که تغییر بیان یکسانی را در حالت سلولی مورد نظر دارند (با هم ، افزایش و یا کاهش می یابند). در شکل ۴ آنالیز خوشه بندی ، پروتئین ها را از نظر بیان به ۲ خوشه اصلی تقسیم نموده است . خوشه سمت چپ دسته پروتئینی را نشان می دهد که دارای بیان کمتری در گروه ۱ و۲ نسبت به گروه ۳ و ۴ بوده اند و در خوشه سمت راست این نسبت تغییر می یابد و گروه های ۱و۲ بیان بیشتر از گروه های ۳و۴ دارند. باید به این نکته نیز توجه داشت که هر چه تعداد شرایط آزمایش افزایش می یابد تعداد زیر خوشه ها نیز افزایش می یابد زیرا با الگوهای بیان متفاوتی از پروتئین ها در بین گروهها روبرو خواهیم شد. ابزار های بسیاری برای تحلیل مجموعه های پروتئین ها وجود دارند، اما اغلب آن ها نتایج حاصله را در یک تصویر ساده و شفاف و قابل تفسیر ارائه نمی دهند. هدف از خوشه بندی در تحقیق بررسی خوشه بندی بر اساس میزان بیان بود. خوشه بندی بر اساس بیان ، نشان دهنده وجود پروتئین هایی است که می توانند وابستگی بیولوژیک نسبت به هم داشته باشند ، به طوری که افزایش بیان همزمان چندین پروتئین باید در جبران کاهش بیان یا عدم بیان پروتئین های دیگر باشد. در واقع ، این عمل یک مکانیسم برقراری هموستازی سلولی است که با خاموشی یک فرآیند ، فرآیند جدیدی جایگزین آن شده و شروع به بیان جدید یا افزایش بیان پروتئین ها می کند. مطالعات پیشین نیز که با کمک این نرم افزار به آنالیز ژل های دو بعدی پرداخته اند نیز نتایج مشابهی را نشان می دهند و بیان می کنند که پروفایل بیانی مشابه اعضای یک خوشه دارای تفسیر بیولوژیکی منطقی است (۱۳–۱۵ و۲۴).

در نرم افزار های آنالیز ژل آنالیز آماری چند متغیره دیگری به نام آنالیز مولفه اصلی (PCA)برای تایید آنالیز خوشه بندی انجام می شود که در شکل۵ نشان داده شده است . اگر ما در جایی مجبور هستیم مهم ترین متغیر را یا یک تعداد محدودی از متغیر ها را در یک مجموعه انتخاب کنیم از آنالیز اجزای اصلی کمک می گیریم . این روش یکی از الگوهای تشخیص و شناسایی (تشخیص هویت)در یک مجموعه اطلاعات است. در این روش اطلاعات را بر اساس شباهت ها و تفاوت هایشان بیان می کنند. از آن جا که در اطلاعات از ابعاد بالا، نقشه و طرح خاصی را به سختی می توان در داده ها پیدا کرد در حقیقت آنالیز اجزای اصلی ارتباط بین داده ها را کشف می کند و در جایی که نعمت نمایش گرافیکی در دسترس نیست، آنالیز اجزای اصلی یک ابزار نیرومند برای آنالیز اطلاعات است که کاهش تعداد ابعاد بدون از دست دادن مقدار

[Downloaded from sjimu.medilam.ac.ir on 2026-11-11]

زیادی از اطلاعات انجام می پذیرد و هدف آن خلاصه کردن داده ها است و به عنوان یک وسیله دسته بندی کننده اطلاعات مورد توجه نیست. در شکل ۵ جایگاه پروتئین های هر دو گروه را نشان می دهد و بیانگر این است پروتئین هایی که متمایل به یک سمت هستند خوشه های مشترک زیادی با هم دارند . از طرفی بیانگر این است که هیچ outlierی در داده وجود ندارد و داده بر طبق شرایط آزمایش خوشه‌بندی شده است(۱۴). PCA کاربردهای زیادی در داده های چند بعدی پروتئومیکی دارد. نتایج آنالیز PCA پراکندکی اعضا هر خوشه در اطراف یک بردار را نشان می دهد و بیانگر این است که نقاط نزدیک به یکدیگر متعلق به یک کلاس هستند که در پروتئومیکس تنوع را در بین مجموعه داده می تواند دسته بندی کند و اعضائ هر دسته را نیز مشخص نماید . از طرفی اگر گونه های متفاوتی از داده مورد استفاده قرار گیرد می تواند خوشه های هر گونه را نیز مشخص نماید و نشان دهد که همه نمونه های یک گونه در اطراف یک بردار واضح قرار می گیرند و همه خوشه ها به هم نزدیک هستند. برای نمونه در مطالعه پروتئومیکی از PCA جهت ارزیابی مجموعه داده ها استفاده شده بود ، همه نمونه های N.meningitidis به طور واضحی در اطراف یک بردار قرار گرفتند و نمونه های که توسط خوشه بندی به روش شبکه عصبی مصنوعی خوشه بندی نشده بودند و در کلاس ها قرار نگرفته بودند در فاصله ای از آنها که خوشه بندی شده بودند قرار گرفتند و نشان می داد که با توجه به ماهیت داده ها که از دو گونه متفاوت بوده است PCA می تواند خوشه های تشکیل شده از روش های مختلف خوشه بندی را مورد ارزیابی قرار دهد وoutliers موجود در مجموعه داده را تعیین نماید(۲۵).نمایش گرافیکی برای نمونه های مورد آزمایش را نشان می دهد. نمایش گرافیکی Power در شکل ۶ حکایت از ۷۲/۲٪ از نقاط استفاده شده در آزمایش است که Power ۸۰٪ یا بیشتر را می‌دهد و چگونگی ارتباط با تعداد تکرار را بیان می‌نماید. برای این Power در این نمونه ها حدود ۳۰ تکرار لازم است.

در نهایت نتیجه این مطالعه نشان داد که استفاده از نرم افزار آنالیز ژل می تواند تفاوت معنی داری آماری در بیان ژن در تمایز سلول بنیادی به آستروسیت و مراحل بلوغ و پیری سلول آستروسیت را که در سطح ملکولی با کاهش و افزایش بیان پروتئین ها و حضور یا ناپدید شدن پروتئین ها خود را نشان می دهد تعیین نماید. این تغییرات را در سطح ملکولی با آنالیزهای آماری چند متغییره ای نظیر کلاسترینگ یا آنالیز همبستگی و آنالیز مولفه اصلی که توسط نرم افزار انجام شد به خوبی آشکار گردید و گروه های پروتئینی که دارای تغییر بیان بودند مشخص شد. برای شناسایی مکانیسم ملکولی نیاز به مشخص شدن نوع پروتئین هایی که در خوشه بندی ها دسته بندی شده بودند وجود دارد که باید توسط دستگاه اسپکتروفوتومتری جرمی تعیین هویت شوند.

سپاسگزاری

این مطالعه حاصل طرح تحقیقاتی با کد -1-1391 150-9123است که توسط کمیته پژوهشی دانشجویان دانشگاه علوم پزشکی شهید بهشتی تصویب گردیده است .از مرکز تحقیقات پروتئومیکس دانشگاه علوم پزشکی شهید بهشتی که در اجرای پروژه همکاری داشتند تشکر می گردد.

References

1- Aittokallio T, Salmi J, Nyman TA, Nevalainen OS. Geometrical distortions in two-dimensional gels: applicable correction method. J Chromatogr B 2005; 815:25-37.
2- Alban A, Davis SO, Bjorkesten L, Andersson C. A novel experimental design for comparative two-dimensional gel analysis: two dimensional difference gel electrophoresis incorporating a pooled internal standard. Proteomics 2003; 3:36-44.
3-Alter O, Brown PO, Botstein D. Singular value decomposition for genome-wide expression data processing and modeling. Proc Natl Acad Sci USA 2000;97:10101-6.
4- Wulfkuhle JD, Liotta LA, Petricoin EF. Proteomic applications for the early

[Downloaded from sjimu.medilam.ac.ir on 2026-11-11]

detection of cancer. Nature Rev 2003; 3:267-74.
5- Kolch W, Mischak H, Pitt AR. The molecular make-up of a tumour: proteomics in cancer research, Clin Sci 2005; 108: 369-83.
6- Patterson SD and Aebersold RH. Proteomics: the first decade and beyond. Nature Genet 2003;33: 311-23.
7- Reif DM, White BC and Moore JH. Integrated analysis of genetic, genomi and proteomic data. Proteomics 2004;1: 67-75.
8- Agaton C, Uhlén M, Hober S., Genome-based proteomics. Electrophoresis 2004:25: 1280-8.
9- Wolfgang Urfer, Marco Grzegorczyk, Klaus Jung , study, a protocol on image acquisition has also been included. Proteomics 2006; 6 : 48-55.
10- Harrington PB, Vieira NE, Jimmy E, Jyh KN, Roberto R, Alfred LY. Analysis of variance–principal component analysis: A soft tool for proteomic discovery. Analytica Chimica Acta 2005; 544: 118-27.
11- Pedreschi R, Hertog ML, Carpentier SC. Treatment of missing values for multivariate statistical analysis of gel-based proteomics data. Proteomics 2008; 8: 1371-83.
12-Bedon F, Villar E, Vincent D, Dupuy JW, Lomenech AM, Mabialangoma A, et al. Proteomic plasticity of two Eucalyptus genotypes under contrasted water regimes in the field. Plant Cell Environ 2012;35:790-805.
13- Bièche C, de Lamballerie M, Chevret D, Federighi M, Tresse O. Dynamic proteome changes in Campylobacter jejuni 81-176 after high pressure shock and subsequent recovery. J Proteomics 20122;75:1144-56.
14- Brunelli L, Campagna R, Airoldi L, Cauli O, Llansola M, Boix J, et al. Exploratory investigation on nitro- and phospho-proteome cerebellum changes in hyperammonemia and hepatic encephalopathy rat models. Metab Brain Dis 2012;27:37-49.
15-Chan K, Casjens S, Parveen N. Detection of established virulence genes and plasmids to differentiate Borrelia burgdorferi strains. Infect Immun 2012;80:1519-29.
16- Meunier B, Dumas E, Piec I, Bechet D, Hebraud M, Hocquette JF. Assessment of hierarchical clustering methodologies for proteomic data mining. J Proteome Res 2007; 6:358-66.
17- Appel R, Hochstrasser D, Roch C, Funk M, Muller AF, Pellegrini C. Automatic classification of two-dimensional gel electrophoresis pictures by heuristic clustering analysis: a step toward achine learning. Electrophoresis 1988; 9:136-42.
18- Wall ME, Rechtsteiner A, Rocha LM. Singular value decomposition and principal component analysis. In: Berrar DP, Dubitzky W, Granzow M, editors. A practical approach to microarray data analysis. Kluwer:Norwell; 2003. P. 91–109.
19- Weisstein EW. Bonferroni correction. 2007 From Math World- a Wolfram web resource.
http://mathworld.wolfram.com/BonferroniCorrection.
20- Wilkins MR, Appel RD, Van Eyk JE, Chung MC, Görg A, Hecker M, et al. Guidelines for the next 10 years of proteomics. Proteomics 2006;6:4–8.
21- Arora PS, Yamagiwa H, Srivastava A, Bolander ME, Sarkar G. Comparative evaluation of two two-dimensional gel electrophoresis image analysis software applications using synovial fluids from patients with joint disease. J Orthop Sci 2005; 10: 160–6.
22-Berth M, Moser FM, Kolbe M, Bernhardt J. The state of the art in the analysis of two-dimensional gel electrophoresis images. Appl Microbiol Biotechnol 2007;76 : 1223-43.
23-Bandow JE, Baker JD, Berth M. Improved image analysis workflow for 2-D gels enables large-scale 2-D gel-based proteomics studies--COPD biomarker discovery study. Proteomics 2008;8 : 3030–41.
24-O'Gorman, M., Beauvallet, C., An investigation into Crohn's disease using the progenesis same spots analysis platform. Nonlinear Dynamics, Newcastle-upon-Tyne, UK, Institut National de Recherche Agronomique (INRA), Jouy-en-Josas, France, Hopital Saint-Antoine, Paris, France.
25- Lancashire L, Schmid O, Shah H, Ball G. Classification of bacterial species from proteomic data using combinatorial approaches incorporating artificial neural networks, cluster analysis and principal

[Downloaded from sjimu.medilam.ac.ir on 2026-11-11]

components analysis. Bioinformatics 2005; 21:2191-9.

Data Analysis of Two Dimensional Electrophoresis Gels by Multivariate Statistical Methods

Zali H1, Rezaei Tavirani M2, Sayedkhani NahalA3, Shahriari Noor M4, Ilghari N4*

(Received: 1 Jan. 2013 Accepted: 27 Feb. 2013)

Abstract

Introduction: Proteomic analysis of massive data sets with having high variable need multivariate methods to provide the simultaneous analysis of multiple variables. In this study, the process of cellular differentiation and aging of stem cells to astrocytes studied by proteomics and gel analysis software and statistical applications were considered.

Materials and Methods: Proteome of four groups (stem cells, young astrocytes, moderately differentiated astrocytes and old astrocytes) were analyzed by the software of Progenesis Same Spots. Cluster analysis, principal component analysis and analysis of power were used in the experimental groups.

Findings: In bioinformatics and statistical analysis of two-dimensional gel electrophoresis technique were detected 940 protein spots with significant expression changes ($p < 0.05$) in four groups that comparisons between groups suggest that the expression of new proteins and the silencing of certain proteins in the signaling pathway of cell differentiation and senescence.

Discussion & Conclusion: Clustering analysis of the expression of proteins can be divided into two main clusters indicate that there are clusters of proteins with similar expression which these proteins can provide similar performance in terms of testing or indicate its presence in the same biological pathway. PCA analysis confirmed the clustering results showed that the protein has been classified according to the test conditions. Finally, the results indicated that by using of the statistical analysis software it is possible to quickly and easily demonstrate that a significant expression changes has been induced in the process of differentiation and senescence on proteome level.

Keywords: Two-dimensional gel electrophoresis, Multivariate statistical technique, Protein

1.Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
2.Proteomics Research Center, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
3. Dept of Clinical Biochemistry, Faculty of Medicine, Ilam University of Medical Sciences, Ilam, Iran
4. Dept of Microbiology Science and Research Branch, Islamic Azad University, Gilan, Iran
**(corresponding author)*

[Downloaded from sjimu.medilam.ac.ir on 2026-11-11]