



طبقه‌بندی مولکولی جدید برای هدایت استراتژی‌های پزشکی دقیق در سندرم شوگرن اولیه

خلاصه

در حال حاضر هیچ درمان تایید شده‌ای برای سندرم شوگرن اولیه وجود ندارد. این بیماری عمدتاً زنان بالغ را تحت تاثیر قرار می‌دهد. مشکل توسعه درمان‌های موثر تا حدودی به دلیل ناهمگونی در تظاهرات بالینی و پاتوفیزیولوژی بیماری است. یافتن خصوصیات مولکولی مشترک در میان زیر گروه‌های بیماران می‌تواند باعث بهبود درک ما از علت بیماری شده و توسعه درمان‌های هدفمند را تسهیل کند. در این مطالعه کوهورت مقطعی، یک طبقه‌بندی مولکولی برای بیماران مبتلا به سندرم شوگرن را گزارش می‌کنیم. این طبقه‌بندی بر اساس پروفایل مولتی امیکس نمونه‌های خون از یک گروه اروپایی متشکل از بیش از ۳۰۰ بیمار و تعداد یکسانی از افراد سالم با پراکندگی سن و جنس مشابه، ایجاد شد. با استفاده از داده‌های رونویسی، ژنومی، اپی ژنتیکی، بیان سیتوکین و فلوسایتومتری، همراه با پارامترهای بالینی، ما چهار گروه از بیماران را با الگوهای مشخصی از اختلالات ایمنی شناسایی کردیم. بیومارکرهایی که ما شناسایی کردیم می‌توانند در آینده توسط یادگیری ماشین برای دسته‌بندی بیماران در زیر گروه‌های مختلف استفاده شوند و امکان ارزیابی مجدد پاسخ‌ها به درمان‌ها در کارآزمایی‌های بالینی را فراهم کنند.

مقدمه

سندرم شوگرن اولیه (pSS) یک بیماری خودایمنی سیستمیک مزمن، ناتوان کننده و پیچیده است که بیشتر، زنان بالغ را تحت تاثیر قرار می‌دهد



فاطمه محمدی پور^۱

۱- کارشناسی بیوتکنولوژی، دانشگاه الزهراء، تهران، ایران
پژوهشگر مرکز تحقیقات پزشکی شخصی آمیتیس ژن



هستند، بنابراین تغییر قابل توجهی در سیر طبیعی خود در مورد شروع، انتشار و شعله‌ور شدن آن‌ها نشان می‌دهند.

مطالعه‌ی حاضر به منظور ایجاد یک طبقه‌بندی مولکولی دقیق از بیماران مبتلا به pSS به دسته‌های همگن‌تر (صرف نظر از فنوتیپ بیماری، فعالیت یا درمان آن‌ها) انجام شد. در اینجا ما در مورد پروفایل مولکولی یکپارچه ۳۰۴ بیمار pSS در مقایسه با ۳۳۰ داوطلب سالم همسان (HV) به دست آمده توسط داده‌های مولتی امیکس با توان بالا در پروژه PRECISESADS IMI JU را گزارش می‌کنیم (ژنتیک، اپی ژنومیک، ترانسکریپتومیک، همراه با داده‌های فلوسایتومتری، سیتوکین‌های مالتی پلکس، و همچنین سرولوژی کلاسیک و داده‌های بالینی). در این مطالعه چهار گروه از بیماران را با الگوهای مشخصی از اختلالات سیستم ایمنی شناسایی شد. خوشه ۱ C3، C1 و C4 IFN بالایی را نشان می‌دهد که منعکس‌کننده درگیری پاتولوژیک مسیر IFN با غنی‌سازی‌های مختلف ژن IFN نوع I و II ژن IFN است. C1 دارای قوی‌ترین امضای IFN با غنی‌سازی ژن نوع I و نوع II در مقایسه با C3 (متوسط) و C4 (پایین تر) است. C4 دارای غنی‌سازی ژن نوع II بیشتر از نوع I و معادل C3 است، در حالی که C3 دارای ترکیب مخالف است. C2 دارای یک بیان ضعیف IFN نوع I و نوع II بدون مشخصات قابل تشخیص آشکار دیگری نسبت به HV است. ما بیشتر C3، C1 و C4 را با استفاده از داده‌های مولتی امیکس و بالینی مشخص کردیم. بیماران C1 شیوع بالایی از SNP‌ها را نشان می‌دهند، بیماران C3 درگیری جزء سلول B برجسته‌تر از سایر خوشه‌ها و عمدتاً افزایش فراوانی سلول‌های B در خون هستند.

در مقابل، بیماران C4 دارای خصوصیات التهابی هستند که توسط مونوسیت‌ها و نوتروفیل‌ها و وضعیت متیلاسیون نابجا هدایت می‌شود. الگوریتم‌های مشتق شده از یادگیری ماشینی، ۴ خوشه را بر اساس بیومارکرهای مختلف متمایز می‌کنند که می‌توانند به راحتی در یک مدل ترکیبی برای طبقه‌بندی بیماران در آزمایشات بالینی استفاده شوند. این مدل ترکیبی با استفاده از یک گروه اولیه مستقل از ۳۷

و هنوز درمان خاصی ندارد. اگرچه درگیری غدد بزاقی و اشکی علامت بارز این بیماری است، در طول پیشرفت pSS، اندام‌ها و سیستم‌های مختلفی از جمله مفاصل، ریه‌ها، کلیه‌ها، کبد، سیستم عصبی و اسکلتی عضلانی نیز می‌توانند درگیر شوند. بنابراین، طیف بالینی بیماری شامل غده‌ی درون‌ریز خوش‌خیم تا یک اختلال سیستمیک شدید با ناهمگنی علائم قابل توجه و عوارض پراکنده متغیر است. تشخیص pSS بر اساس ترکیبی از پارامترهای بالینی، سرولوژیکی، بافت‌شناسی و عملکردی است که اغلب تنها در مراحل پایانی بیماری، یعنی زمانی که اختلال عملکرد غدد و علائم به شدت بر کیفیت زندگی کلی بیمار تأثیر می‌گذارد، صورت می‌گیرد. علاوه بر این، یک پنجم بیماران pSS ممکن است درگیری اندام مرکزی با آسیب بالقوه شدید اندام انتهایی را نشان دهند و پنج درصد از بیماران نیز ممکن است به لنفوم غیر هوچکین مبتلا شوند. pSS یکی از محدود بیماری‌های اولیه است که خودایمنی، توسعه سرطان و عفونت‌ها را به هم مرتبط می‌کند و بینش منحصربه‌فردی را در مورد بسیاری از علوم پایه و حوزه‌های پزشکی بالینی ارائه می‌دهد. با این حال، پاتوژنز بیماری مبهم باقی مانده است. به طور خاص، دانش محدود از انواع بیماری pSS موجود مسلماً بزرگترین مانع برای بهبود تشخیص بیماران و شناسایی زیرمجموعه‌های بیماران با توجه به طبقه‌بندی اولیه و درمان شخصی شده است. اخیراً در پروژه PRECISESADS IMI JU نشان داده شد که بیماری‌های خودایمنی سیستمیک طیف متنوع و یک فنوتیپ مولکولی پیچیده یا همپوشانی را با چهار خوشه شناسایی شده نشان می‌دهند که نشان‌دهنده الگوهای «inflammatory»، «lymphoid»، «interferon» و «healthy-like» هستند. که هر کدام شامل تمام تشخیص‌ها می‌شود و با ویژگی‌های ژنتیکی، بالینی، سرولوژیکی و سلولی تعریف می‌شود. بسیاری از آن‌ها ژن‌های مشترک دارند و بیان بیش از حد ژن‌های القایی اینترفرون (IFN) معروف به امضای IFN در بسیاری از این بیماران مشاهده می‌شود. چنین بیماری‌های خودایمنی تحت تأثیر عوامل محیطی متعددی

بیمار pSS تایید شده است. در نتیجه، این کار درک روشنی از ناهمگنی pSS ارائه می‌کند، و امضاهای مرتبط بالینی و ایمنی را برای هدایت استراتژی‌های پزشکی دقیق ارائه می‌کند. تصمیم‌گیری مبنی این طبقه بندی بیماران بلافاصله برای ارزیابی مجدد پاسخ‌های درمانی در کارآزمایی‌های بالینی اعمال می‌شوند.

نتایج

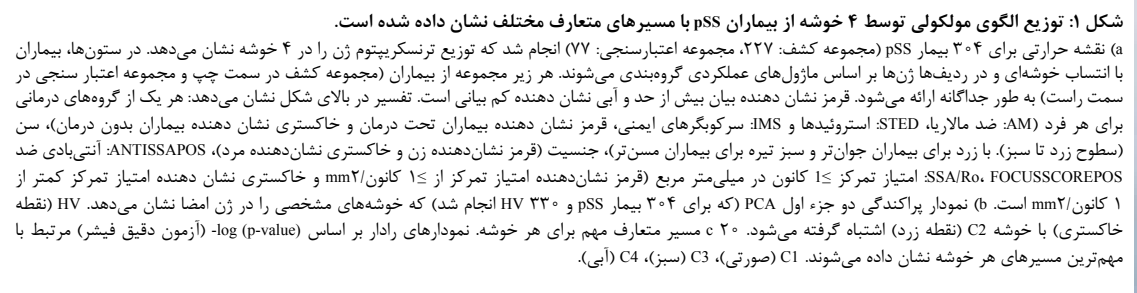
شناسایی چهار خوشه مولکولی عملکردی از بیماران pSS

جمعیت مطالعه اولیه ما که در مطالعه مقطعی PRECISESADS ثبت نام کردند شامل ۳۸۲ بیمار pSS بود. پس از کنترل کیفیت کامل و تایید تشخیص (هر بیمار باید آنتی‌بادی ضد SSA/Ro یا سیالانیت لنفوسیتی کانونی را با امتیاز تمرکز ≥ 1 کانون در میلی‌متر مربع ارائه می‌کرد)، ۷۸ بیمار حذف شدند. ویژگی‌های بیماران در جدول ۱ ارائه شده است. برای خوشه بندی ۳۰۴ نمونه باقیمانده، داده‌های ترنسکریپتومیکس با یک رویکرد قوی نیمه نظارتی که قبلاً برای سرطان پستان اعمال شده بود، آنالیز شد. نمونه‌ها به یک مجموعه کشف و

یک مجموعه اعتبارسنجی تقسیم شدند که ۷۵٪ و ۲۵٪ از قطعات را نشان می‌دهد. مجموعه اکتشاف برای خوشه‌ای از بیماران در چهار گروه، همانطور که در مجموعه اعتبارسنجی تایید شد (شکل ۱a) اجازه داد. هنگامی که دو مکان ادغام شدند، خوشه ۱ (C1) شامل ۱۰۱ بیمار (۳۳.۲٪)، خوشه ۲ (C2) ۷۷ بیمار (۲۵.۳٪)، خوشه ۳ (C3) ۸۸ بیمار (۲۸.۹٪) و خوشه ۴ (C4) ۳۸ نفر (۱۲.۵٪) بود. مرحله نظارت امکان انتخاب زیرمجموعه‌ای از ۲۵۷ ژن برتر را فراهم می‌کند که ۴ خوشه از بیماران را متمایز می‌کند و به سه ماژول تقسیم می‌شود: M.a (۱۰۵) ژن، M.b (۲۰) ژن و M.c (۱۳۲) ژن. آنالیز غنی‌سازی برای تفسیر هر یک از ماژول‌های ژنی مورد استفاده قرار گرفت، که نشان داد M.a در سیگنال دهی IFN، M.b در مسیرهای دودمان لنفاوی و M.c در رونوشت‌های دودمان التهابی و میلوئیدی غنی شده است. C1، و تا حدی کمتر، C3، بیان بیش از حد ماژول ژن M.a را نشان داد، در حالی که C3 بیان بیش از حد M.b را نیز نشان داد و C4 بیان بیش از حد قوی M.c را نشان داد (شکل ۱a). از آنجایی که C2 هیچ الگوی قابل تشخیص آشکاری نداشت، داوطلبان سالم (HV) به فاصله چهار خوشه مولکولی تا مرکزها اختصاص داده شدند (شکل ۱b). هنگامی که در جمعیت بیمار پیش‌بینی شد، HV یک خوشه جداگانه تشکیل نمی‌داد، بلکه عمدتاً با C2 مطابقت داشت (به ترتیب ۰.۵٪، ۹۳٪، ۴٪ و ۲.۵٪ HV با C1، C2، C3 و C4 ادغام شدند). این بدان معنی است که امضای ترنسکریپشنال C2 حداقل در سطح خون متفاوت از HV نیست. جالب توجه است، داده‌های ما با مشاهدات قبلی یک گروه بیمار مشابه سالم که در یک جمعیت تلفیقی از ۷ بیماری خودایمنی مختلف شناسایی شده بود، مطابقت دارد.



جدول ۱ - داوطلبان سالم (HV) و ویژگی‌های بیمار با سندرم شوگرن اولیه (pSS).						
pSS All (N = 304)	pSS Validation (N = 77)	pSS Discovery (N = 227)	HV (N = 330)			
						دموگرافی
304	77	227	330	n		سن
58.401 ± 13.448	58.039 ± 13.554	58.524 ± 13.440	53.294 ± 10.998	Mean ± SD		
304	77	227	330	n		جنسیت
282 (92.76)	71 (92.21)	211 (92.95)	302 (91.52)	n (%)	زن	
292	74	2108	238	n		چاقی (BMI >= 30)
33 (11.30)	3 (4.05)	30 (13.76)	24 (7.27)	n (%)	بله	
304	77	227	330	n		نژاد
2 (0.66)	1 (1.30)	1 (0.44)	2 (0.61)	n (%)	آسیایی	
1 (0.33)	1 (1.30)	-	-	n (%)	سیاه پوست / آفریقایی آمریکایی	
298 (98.03)	74 (96.10)	224 (98.68)	328 (99.39)	n (%)	قفقازی / سفید	
3 (0.99)	1 (1.30)	2 (0.88)	-	n (%)	باقی	
						معیارهای تشخیصی
109	27	82	-	n		امتیاز تمرکز < ۱
97 (88.99)	24 (88.89)	73 (89.02)	-	n (%)	بله	
304	77	227	-	n		ضد SSA مثبت
274 (90.13)	69 (89.61)	205 (90.30)	-	n (%)	بله	
						فعالیت بیماری
302	77	225	-	n		مدت بیماری، سال
10.866 ± 8.101	11.094 ± 9.620	10.788 ± 7.535	-	میانگین ± SD		
286	75	22=11	-	n		فعالیت بیماری (PGA*)
25.465 ± 19.488	24.840 ± 20.984	25.687 ± 18.976	-	میانگین ± SD		
193	60	133	-	n		ESSDAI (**)
4.684 ± 5.388	4.850 ± 5.495	4.609 ± 5.358	-	میانگین ± SD		
150	44	106	-	n		ESSPRI (**)
4.998 ± 2.405	4.568 ± 2.648	5.176 ± 2.286	-	میانگین ± SD		
n: تعداد بیماران با اطلاعات موجود. PGA (*): ارزیابی کلی پزشک. (**) در یک مطالعه فرعی جمع آوری شده است.						





غنی شد. فعال سازی فاکتور (IRF) قابل توجه، C3 و C4 بیشتر با تغییرات در شبکه‌های بیولوژیکی مرتبط با ایمنی تطبیقی مشخص شدند. به طور خاص، فعال شدن قابل توجه مسیرهای متعارف مربوط به فعال سازی سلول B مانند سیگنال دهی گیرنده سلول B، و توسعه سلول B در C3 مشاهده شد. علاوه بر این، آنالیزهای مقایسه‌ای شواهدی برای تنظیم بالا سیگنالینگ IL7 و فعال سازی LXR/RXR در C3 در مقایسه با C1 ارائه کردند.

جالب توجه است، C4 اندوتیپ با بیشترین تعداد DEG در مقایسه با HV با مسیرهای متعارف بسیار ناهمگن بود. آنالیز مسیر ابتکاری فعال شدن مسیرهای مربوط به لنفوسیت‌های T و B را تایید کرد که منعکس کننده فعال سازی Th1 و Th2، سیگنال دهی گیرنده سلول B، همراه با علائم التهابی برجسته است که به ویژه با سیگنال دهی سیتوکین مرتبط است (IL-6 و IL-10 و سیگنالینگ، تولید IL-15، مسیر STAT-3).

آنالیز بیشتر تنظیم کننده بالادستی فعال شدن قابل توجه IFN- α را در هر سه خوشه، و همچنین CpG ODN در C3 و TNF- α ، IFN γ ، LPS، و IL-4 در C4 را پیش‌بینی کرد، که فعالیت سلول B و پاسخ‌های التهابی را به ترتیب در C3، C4 و C3 برجسته‌تر کرد. در حالی که C2 هیچ DEG را در مقایسه با HV نشان نداد، ۱۴ ژن در بیماران C2 به طور متفاوتی برای آنتی‌بادی‌های SSA در مقایسه با HV مثبت بیان شد در حالی که تنها ۲ DEG در بیماران C2 SSA منفی یافت شد. این بیماران C2 مثبت SSA با غنی‌سازی قابل توجهی در ژن‌های مرتبط با IFN در مقایسه با HV از جمله IFI44، IFI44L، IFI6، IFIT1، IFIT3، ISG15، MX1، OAS3، SERPING1 و SIGLEC1 مشخص شدند.

برای مشخص کردن بیشتر تنوع خوشه‌ای بیماران در سطح مولکولی، ما از مجموعه مدولار رونوشت خون که اخیراً بر روی طیف گسترده‌ای از بیماری‌ها و وضعیت‌های پاتولوژیک ایجاد شده است، استفاده کردیم. مورد دوم شامل ۳۸۲ ماژول ترنسکرپتوم بر اساس الگوهای بیان مشترک ژن در ۱۶ بیماری و ۹۸۵ پروفایل رونوشت منحصر به فرد است. باز هم، هیچ تجمعی به طور متفاوت در C2 یافت نشد که پروفایل سالم مانند این بیماران را تأیید کند، در حالی که یک امضای IFN

سپس ارزیابی کردیم که آیا متغیرهای کمکی مانند درمان‌های سیستمیک می‌توانند خوشه‌بندی مبتنی بر ترنسکرپتوم را هدایت کنند یا خیر. در واقع، نیمی از بیماران pSS در زمان ویزیت با داروهای ضد مالاریا، سرکوب‌کننده‌های ایمنی یا استروئیدها تحت درمان قرار گرفتند که از نظر آماری تفاوت معنی‌داری در توزیع بین چهار خوشه وجود داشت (p-value برای ضد مالاریا به ترتیب ۰.۰۰۲ بود، ۰.۰۰۱ < برای سرکوب‌کننده‌های ایمنی و استروئیدها) (جدول ۲). در مقایسه با ۳ خوشه دیگر، نسبت بیشتری از بیماران تحت درمان با داروهای ضد مالاریا در C2 و نسبت بیشتری از بیماران دریافت کننده سرکوب‌کننده‌های ایمنی یا استروئیدها در C4 مشاهده شد. مهمتر از همه، آنالیز حساسیت بیماران تحت درمان در مقابل بیماران درمان نشده در هر خوشه هیچ تاثیری از درمان‌ها بر توزیع خوشه نشان نداد.

آنالیز دقیق مسیر عملکردی خوشه‌های pSS فردی

برای بررسی فرآیندهای مولکولی و عملکرد بیولوژیکی آنها در زیربنای هر یک از خوشه‌های بیماران pSS، امضاهای ژن‌های بیان شده متفاوت (DEG) در مقایسه با HV با استفاده از Limma در ۴ خوشه ارزیابی شد. متعاقباً از آنالیز مسیر هوشمندی (IPA) برای تعیین مسیرهای متعارف بی‌نظم قابل توجهی با مقدار p تنظیم‌شده بنجامینی-هوخرگ $0.05 \leq (FDR)$ و تغییر برابر مطلق $1.5 \geq (FC)$ استفاده شد. در نتیجه، ۲۸۴ درجه در C1، ۳۰۱ درجه در C3 و ۱۶۸۶ درجه در C4 معنی‌دار بود.

از آنجایی که هیچ DEG در C2 در مقایسه با HV مشاهده نشد، تنها C1، C3 و C4 از نظر عملکردی تفسیر شدند. ۲۰ مسیر متعارف مهم در هر امضای DEG ارائه شده‌اند و مسیرهای مربوط به پاسخ‌های ایمنونولوژیکی غنی‌شده‌تر به عنوان نمودارهای راداری در شکل ۱c گزارش شده‌اند. در حالی که هر ۳ خوشه در ژن‌های دخیل در پاسخ‌های ضدویروسی و ضدباکتریایی غنی شده بودند که نشان‌دهنده یک پروفایل فعال‌سازی با واسطه ذاتی است، C1 عمدتاً با مسیرهای مرتبط با IFN از جمله سیگنال‌دهی IFN، نقش گیرنده‌های تشخیص الگو برای باکتری‌ها و ویروس‌ها و تنظیم کننده اینترفرون



جدول ۲ آنالیز توصیفی پارامترهای بالینی توسط خوشه سندرم شوگرن اولیه.							
p-value	C4 (n = 38)	C3 (n = 88)	C2 (n = 77)	C1 (n = 101)			
	38	88	77	101	n		سن و سال
0.10	63.105 ± 14.790	57.250 ± 12.032	58.805 ± 13.688	57.327 ± 13.705	میانگین ± SD		
	38	88	77	101	n		جنسیت
0.70	34 (89.47)	81 (92.05)	71 (92.21)	96 (95.05)	n (%)	زن	
	37	88	76	101	n		سن شروع، سال
0.071	51.739 ± 16.053	47.606 ± 12.687	50.428 ± 14.532	45.663 ± 14.475	میانگین ± SD		
	37	88	76	101	n		مدت بیماری، سال
0.02994	12.625 ± 8.52	10.183 ± 7.210	8.965 ± 7.336	12.247 ± 8.921	میانگین ± SD		
	36	85	71	94	n		فعالیت بیماری (*PGA)
0.092	31.556 ± 20.646	23.212 ± 18.766	22.718 ± 17.698	27.245 ± 20.535	میانگین ± SD		
	27	44	52	70	n		ESSDAI
0.10	6.370 ± 6.828	4.227 ± 4.017	3.731 ± 4.594	5.029 ± 5.959	میانگین ± SD		
	21	30	43	56	n		ESSPRI
0.87	4.937 ± 1.803	5.300 ± 2.703	5.031 ± 2.429	4.833 ± 2.460	میانگین ± SD		
	38	86	77	98	n		آرتروز
0.016	12 (31.58)	20 (23.26)	18 (23.38)	39 (39.80)	n (%)	گذشته	
	5 (13.16)	4 (4.65)	3 (3.90)	2 (2.04)	n (%)	حال حاضر	
	14	21	29	96	n		امتیاز تمرکز < ۱
0.4	12 (85.71)	17 (80.95)	28 (96.55)	39 (40.63)	n (%)	بله	
	38	88	77	101	n		ضد SSA مثبت
<0.001	31 (81.57)	87 (98.86)	56 (72.72)	99 (99.00)	n (%)	بله	
	38	86	77	100	n		
<0.001	11 (28.95)	39 (45.35)	12 (15.58)	61 (61.00)	n (%)	بله	
					n		
					n (%)	بله	
	38	86	73	97	n		هیپرگاماگلوبولینمی
<0.001	3 (7.89)	9 (10.47)	8 (10.96)	23 (23.71)	n (%)	گذشته	
	7 (18.42)	41 (47.67)	10 (13.70)	44 (45.36)	n (%)	حال حاضر	

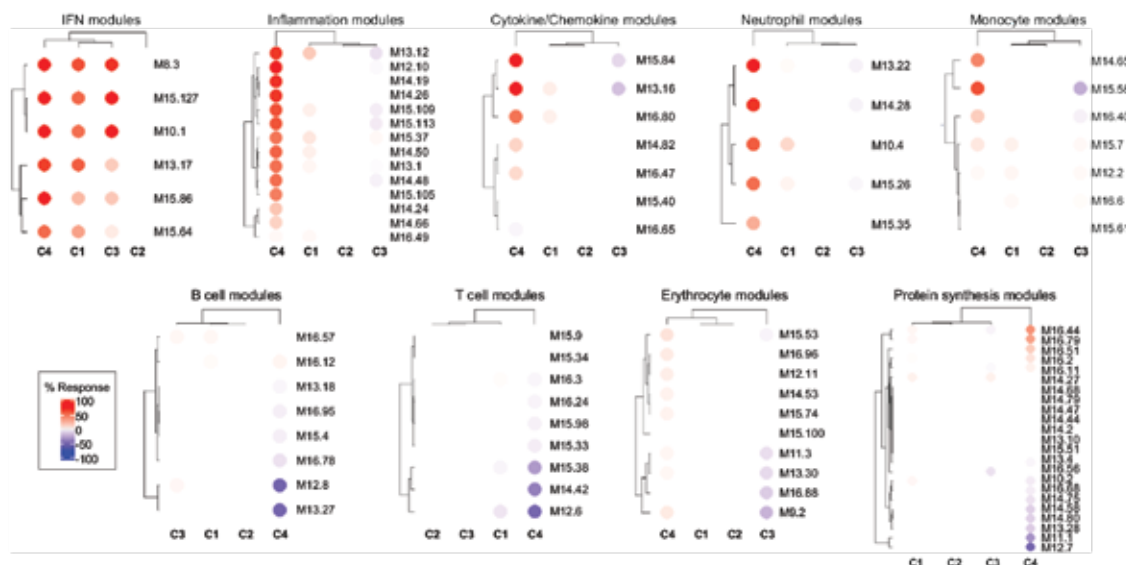


جدول ۲ آنالیز توصیفی پارامترهای بالینی توسط خوشه سندرم شوگرن اولیه.							
	38	87	77	100	n		شاخص‌های التهابی غیرطبیعی
0.003	12 (31.58)	20 (22.99)	13 (16.88)	28 (28.00)	n (%)	گذشته	
	10 (26.32)	22 (25.29)	11 (14.29)	35 (35.00)	n (%)	حال حاضر	
	35	82	74	93	n		کاهش سطح C3
0.8	4 (11.43)	11 (13.41)	5 (6.76)	13 (13.98)	n (%)	گذشته	
	3 (8.57)	5 (6.10)	4 (5.41)	7 (7.53)	n (%)	حال حاضر	
	35	82	74	93	n		کاهش سطح C4
0.10	4 (11.43)	9 (10.98)	3 (4.05)	13 (13.98)	n (%)	گذشته	
	4 (11.43)	3 (3.66)	3 (4.05)	10 (10.75)	n (%)	حال حاضر	
	38	88	77	98	n		کراتینین غیر طبیعی
0.009	2 (5.26)	-	4 (5.19)	10 (10.20)	n (%)	گذشته	
	6 (15.79)	7 (7.95)	2 (2.60)	5 (5.10)	n (%)	حال حاضر	
n	25	56	58	65	n		پروتئینوری
0.093	3 (12.00)	1 (1.79)	2 (3.45)	5 (7.69)	n (%)	در حد متوسط	
	—	3 (5.36)	—	5 (7.69)	n (%)	گذشته	
	38	88	77	101	n		استفاده کنونی از داروهای ضد مالاریا
0.002	15 (39.47)	24 (27.27)	42 (54.55)	33 (32.67)	n (%)	بله	
	38	88	77	101	n		استفاده فعلی از داروهای سرکوب کننده ایمنی
<0.001	15 (39.47)	7 (7.95)	14 (18.18)	17 (16.83)	n (%)	بله	
n	38	88	77	101	n		استفاده کنونی از استروئیدها
<0.001	23 (60.53)	10 (11.36)	14 (18.18)	23 (22.77)	n (%)	بله	
n: تعداد بیماران با اطلاعات موجود، PGA: Physician Global Assessment(*) آزمون‌های آماری انجام شده: آزمون کای دو استقلال برای متغیر طبقه‌بندی و آزمون کروسکال-والیس برای متغیر ادامه.							

امضاهای IFN

مطابق با ادبیات، غنی ترین مسیری که به طور قابل توجهی تایید شده است که در هر سه خوشه تنظیم شده است، مسیر سیگنالینگ IFN بود (شکل ۲). در SLE، چپچه و همکاران. قبلاً سه ماژول مشروح M3.4، M1.2 و IFN (M5.12) را از داده های رونوشت خون محیطی شناسایی کرده اند که برای هر ماژول یک آستانه فعال سازی متمایز وجود دارد. ژن های درون ماژول M1.2 توسط IFN α القا می شوند، در حالی که سایر ژن های هر دو M1.2 و M3.4 توسط IFN β تنظیم می شوند، که مربوط به امضای IFN نوع I است. ژن های M5.12 به تنهایی توسط IFN α و IFN β ضعیف القا می شوند، اما توسط IFN γ مشخص کننده امضای IFN نوع II تنظیم می شوند. علاوه بر این، ترنسکریپتوم متعلق به M3.4 و M5.12 تنها به طور کامل توسط ترکیبی از IFN های نوع I و نوع II القا شدند. کپرو و همکاران مشاهدات مشابهی انجام داد و ژن هایی را شناسایی کرد که ترجیحاً توسط IFN α یا IFN γ 12 القا می شوند. سپس امتیازهای Z مختلف بر این اساس برای مشخص کردن بیشتر امضای IFN مشاهده شده در خوشه های مختلف محاسبه شد (شکل ۳). تمام IFN z-نمرات تا حدی در C2 در مقایسه با HV افزایش

تنظیم شده بالا در C1، C3 و C4 غالب بود (شکل ۲). در C4، ماژول های القا شده شامل ژن های مرتبط با التهاب و نوتروفیل ها هستند. به عنوان بالاترین فنوتیپ التهابی، C4 با هیپرستیکو کینمی/هیپر کمو کینمی مشاهده شده در ماژول ها (M13.16، M15.84، M16.80) همراه است که با تنظیم مثبت ماژول مرتبط با TNF (M16.47) و کاهش میزان ماژول مرتبط با TGF β (M16.65) (شکل ۲). برخی از ماژول ها کمتر بیان شده بودند، مانند آن هایی که با هر دو سنتز پروتئین (M12.7، M11.1، M13.28) (M14.80)، سلول های (M13.27، M12.8) B و سلول های (M15) T مرتبط هستند. ژن هایی که عمدتاً در C1 بیان می شوند نیز در پاسخ های التهابی و نوتروفیل ها (A33، A35)، به موازات امضای سلول های B و T تنظیم شده پایین دخیل بودند. علاوه بر این، ماژول های فرعی مجزا که در جهت های مخالف بیان می شوند، امکان تمایز عملکردی C1 و C3 را فراهم می کنند. بیماران از C3 بیان کمتر قابل توجهی از ماژول های مربوط به گلبول های قرمز (M9.2، M11.3، A37) و سیتوکین ها / کموکین ها (M15.84، M13.16، A35) و افزایش بیان در برخی از سلول های B را نشان دادند.



شکل ۲: الگوهای فراوانی ماژول های مختلف، چهار خوشه pSS را متمایز می کند. هر نقشه حرارتی که با بسته BloodGen3Module R به دست آمد، یکی از مهم ترین الگوهای متمایز کننده چهار خوشه ۳۰۴ بیمار (C1: 101، C2: 77، C3: 88 و C4: 38) pSS را در مقایسه با ۳۳ داوطلب سالم (HV) نشان می دهد. این الگوها با ماژول های مرتبط با IFN، نوتروفیل ها، التهاب، سیتوکین ها/کموکین ها، سنتز پروتئین، گلبول های قرمز، مونوسیت ها، سلول های B و سلول های T مطابقت دارند. ستون های این نقشه حرارتی مربوط به خوشه ها هستند. هر ردیف مربوط به یکی از ماژول های مرتبط با الگو است. برای هر ماژول، درصد افزایش ژن (از ۰ تا ۱۰۰) و ژن کاهش یافته (از ۰ تا ۱۰۰) محاسبه شد. یک نقطه قرمز روی نقشه حرارتی نشان دهنده افزایش فراوانی رونوشت های متشکل از یک ماژول معین برای یک خوشه معین است. یک نقطه آبی نشان دهنده کاهش فراوانی رونوشت ها است. عدم وجود رنگ نشان دهنده عدم تغییر است.

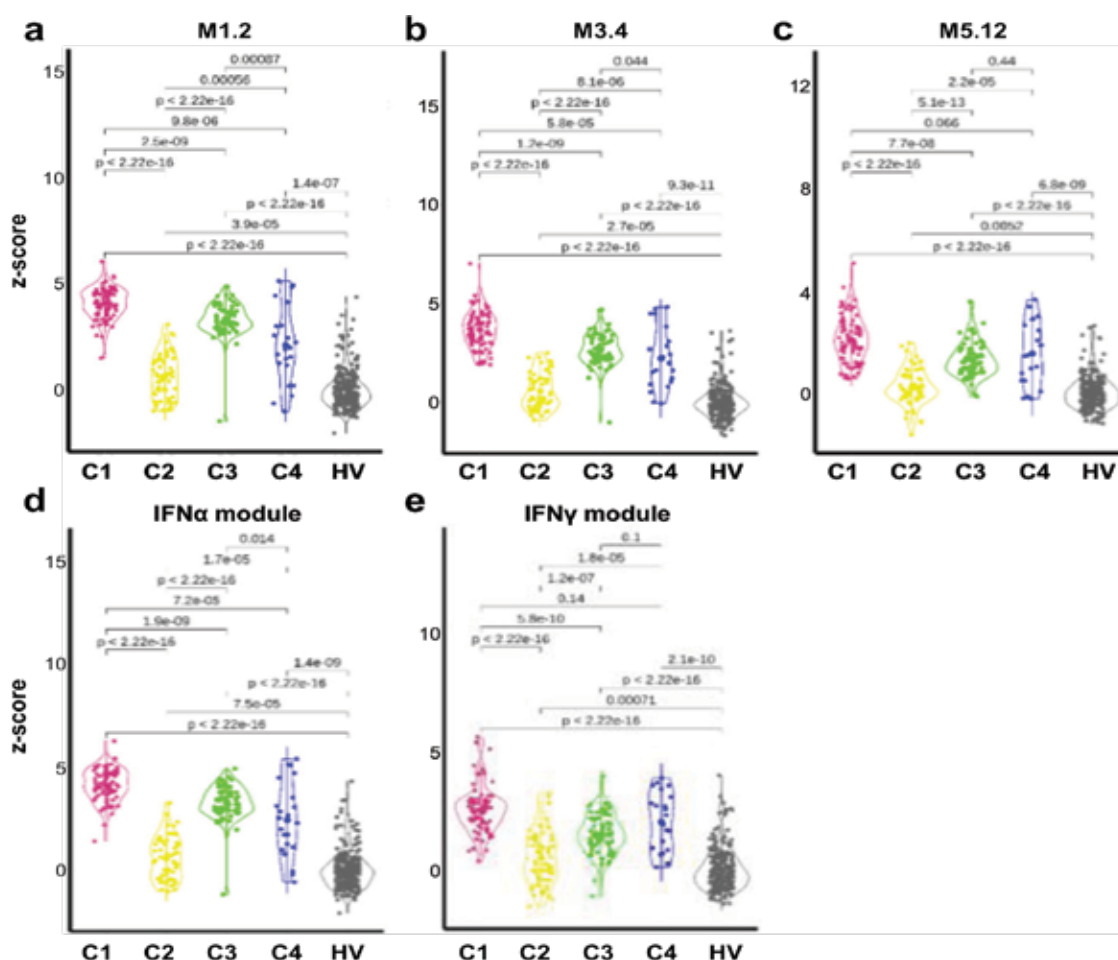


تحلیل شده (pSS ۳۰۴ و HV ۳۳۰)، ما به طور واضح (با سیگنال‌های سطح معنی‌داری گسترده ژنوم 8×10^{-5}) ۳۵ پلی مورفسم تک نوکلئوتیدی (SNPs) در C1 در مقایسه با تنها شش در C3 شناسایی کردیم. و یکی در C4 (شکل ۴a). جالب توجه است که هیچ غنی‌سازی قابل توجهی در C2 یافت نشد. 35 SNP ارزیابی شده در C1 در ژن‌های مرتبط با سیستم ایمنی (HLA-DQB1، HLA-DQA1، HLA-DRA، HLA-C، HLA-G انتقال سیگنال (NOTCH4)، زیست‌شناسی رشدی (POU5F1)، یافت می‌شوند. بیان ژن (DDX39B) یا چرخه سلولی (TUBB). وجود چنین پیوندهای ژنتیکی قابل توجهی قبلاً در خوشه‌هایی از بیماران بیماری

یافت. در راستای سیگنال قوی مشاهده شده، بیماران C1 بالاترین نمرات نوع I و II را داشتند. جالب توجه است که C3 امتیاز IFN نوع I بالاتری نسبت به C4 داشت اما این ۲ خوشه برای نمره IFN نوع II تفاوتی نداشتند. آنالیز بالادست IFN γ DEG C4 را به عنوان یک تنظیم کننده مهم پیش بینی کرد که نشان می‌دهد فعال سازی IFN نوع II در C4 برجسته بود.

آنالیز مطالعه ارتباط ژنومی

ما بررسی کردیم که آیا خوشه‌ها تفاوت‌هایی را در مشارکت ژنتیکی آل‌های خطر شناخته شده با pSS نشان می‌دهند. حتی در گروه متوسط از بیماران تجزیه و



شکل ۳: ۴ خوشه pSS امضای معمولی IFN را مطابق با امتیازهای z مدولار IFN نشان می‌دهند. آنالیز امتیاز IFN برای ۳۰۴ بیمار pSS و ۳۳۰ داوطلب سالم (HV) انجام شد. تقسیم مجدد نمونه‌ها از ۴ خوشه pSS با توجه به مشخص‌ترین امتیازات z ماژول IFN نشان داده شده است. ژن‌های (a) M1.2 توسط IFN α القا می‌شوند، در حالی که ژن‌های (b) M3.4 و M5.12 (ZBP1، IFIH1، EIF2AK2، PARP9 و M1.2) توسط IFN γ تنظیم می‌شوند. علاوه بر این، رونوشت‌های متعلق به M3.4 و M5.12 تنها به طور کامل توسط ترکیبی از IFNs نوع I و نوع II القا شدند. ماژول‌های دیگر ژن‌هایی را که ترجیحاً توسط IFN α (IFIT1، IFI44) یا IFN γ (IRF1، GBP1 یا EIF2AK2) (d) و IFN α (SERPING1 و IFN γ) القا می‌شوند شناسایی کردند (e). نتایج آزمون مجموع رتبه ویلکاکسون جفتی دو دم نشان داده شده است. نمودارها میانه را با نوارهای خط نشان می‌دهند که محدوده بین چارکی را نشان می‌دهد.

لوکوس NOTCH4/C6orf10 دارای پنج SNP مرتبط (rs3130347، rs204991، rs3132935، rs7751896، rs9268220) و همچنین IER3 | لوکوس DDR1 (rs3094122، rs6911628، rs3094112، rs2517576، rs3095151) را ارائه کرد.

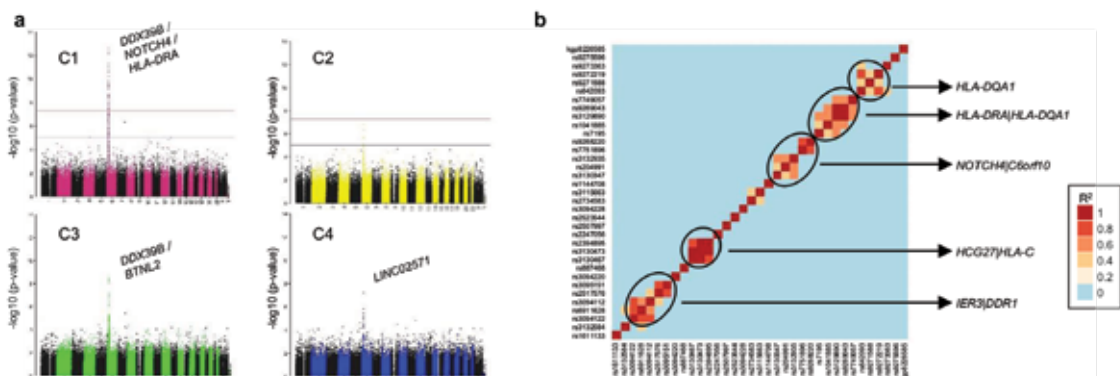
آنالیز متیلاسیون

آنالیز متیلاسیون با بنجامینی هوچبرگ $FDR < 0.1$ و ΔBeta مطلق > 0.075 انجام شد. تنها دو موقعیت متیله متفاوت (DMPs) مربوط به دو ژن در C2 یافت شد. آن DMP ها با ۳ خوشه دیگر مشترک بودند (شکل ۵a) و در ساحل ۱۵۰۰ TSS ژن NLRC5 و در 5'UTR ژن MX1، دو ژن درگیر در امضای IFN، قرار داشتند. NLRC5 از طریق مهار فعال شدن NF-kappa-B و تنظیم منفی مسیرهای سیگنالینگ نوع IFN نقشه در پاسخ سیتوکین و ایمنی ضد ویروسی ایفا می کند. MX1 یک GTPase دینامیک ناشی از IFN را با فعالیت ضد ویروسی کم می کند که به عنوان یک نشانگر زیستی قابل اجرا بالینی برای شناسایی IFN نوع I سیستم در pSS19 پیشنهاد شده است.

۱۴۵ DMP مربوط به ۸۷ ژن و ۹۶ DMP مربوط به ۵۶ ژن به ترتیب در C1 و C3 یافت شد، در حالی که وضعیت متیلاسیون نابجا با ۸۴۴۵ DMP مربوط به ۳۶۳۶ ژن C4 را مشخص می کند (شکل ۵a). به منظور آزمایش اینکه آیا نقص متیلاسیون در C4 با درمان استروئیدی مرتبط است، ۹ بیمار درمان نشده را با

خودایمنی سیستمیک که مسیر بیماری مولکولی آنها مسیر IFN نوع I است، مشاهده شده است. علاوه بر این، ارتباط قوی SNP ها با ژن های HLA کلاس II در بیماران SLE با سطح بالایی از آنتی بادی ها گزارش شد. یک (rs2734583) SNP با C1 و C3 مشترک بود و با ژن DDX39 مرتبط است. شایان ذکر است، DDX39B پروتئین کدگذاری شده توسط این ژن، برای جلوگیری از تشکیل dsRNA در طول عفونت ویروس آنفلوآنزای A مورد نیاز است و در نتیجه از فعال شدن سیستم IFN نوع I جلوگیری می کند. پنج SNP دیگر در C3 عبارتند از، (2 SNP)، HLA-DQA، HLA-DRA، BTNL2 و HCG23. تنها (rs2247056) SNP که در C4 یافت شد، همچنین با C1 مشترک است، در اینترون ۱ ژن LINC02571 قرار دارد و قبلاً با خطر ابتلا به SLE مرتبط بود.

عدم تعادل پیوندی یک ارتباط غیر تصادفی از آلل ها در جایگاه های مختلف در یک جمعیت معین است. هنگام آنالیز عدم تعادل پیوندی (شکل ۴b) در ۳۵ جایگاه های SNP شناسایی شده در C1 و واقع در کروموزوم ۶ (از پایه ۲۹۸۰۹۳۶۲ تا ۳۲۶۸۱۶۳۱)، سه SNP به شدت در مکان HLA-DQA1 (rs919,29,29,29,2927rs) مرتبط بودند. پنج SNP در لوکوس HLA-DQA1 | HLA-DRA (rs7195، rs1041885، rs3129890، rs9269043، rs7749057) و سه SNP در لوکوس HLA-C (rs3130473، rs2394895 و rs3130467). دو منطقه دیگر حاوی SNP های مرتبط قوی هستند.

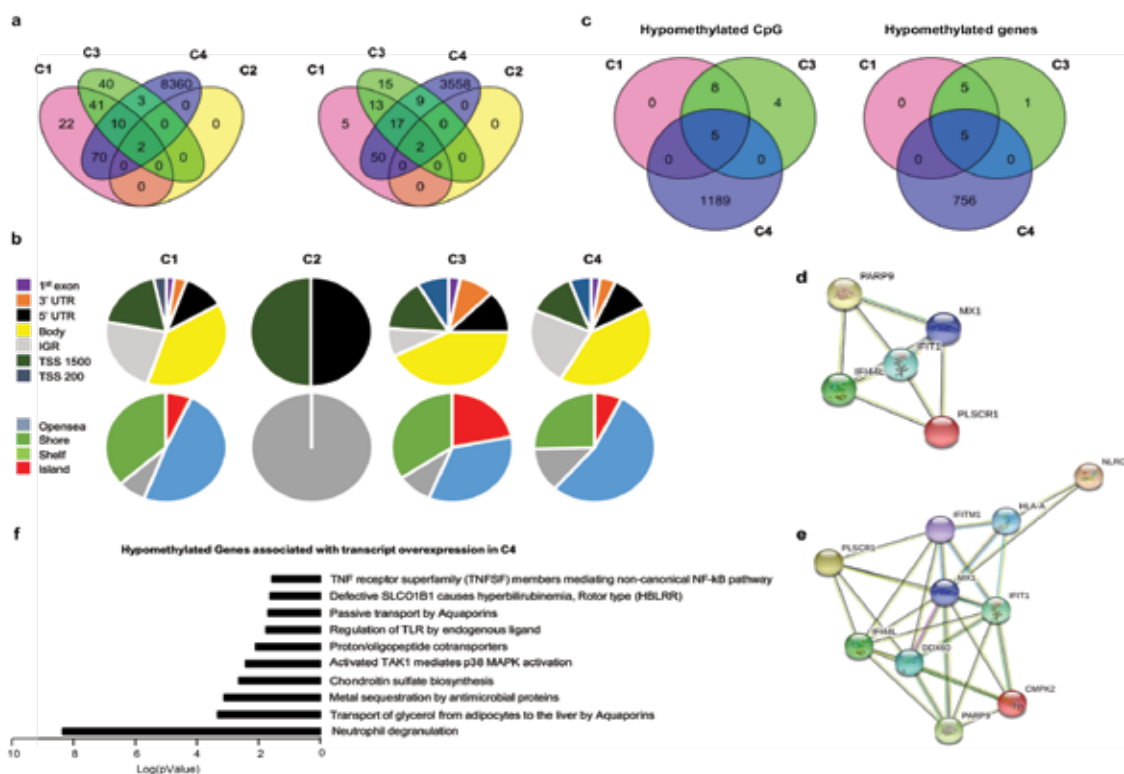


شکل ۴: تجزیه و تحلیل ارتباط ژنومی خوشه ای (GWAS). آنالیز GWAS با استفاده از Plink، یک مجموعه ابزار آنالیز انجمن کل ژنوم منبع باز، با استفاده از رگرسیون لجستیک برای C1: 101، C2: 77، C3: 88 و C4: 38 و 304 pSS (۳۳ داوطلب سالم (HV) و هر خوشه با HV مقایسه شد. یک نمودار منهن برای هر خوشه نشان داده شده است. b آنالیز عدم تعادل پیوند در جایگاه های ۳۵ SNP شناسایی شده در C1 و واقع در کروموزوم ۶ از پایه ۲۹۸۰۹۳۶۲ تا ۳۲۶۸۱۶۳۱. ضریب همبستگی R^2 و نقشه حرارتی عدم تعادل پیوند به ترتیب با بسته بندی Plink و oncofunco R به دست آمدند. قوی ترین ارتباط بین SNP ها مشروح شده است.



تا ۲ کیلوبایت از جزیره CpG)، قفسه ها (مناطق از ۲ تا ۴ کیلوبایت از جزیره CpG) و دریای آزاد (بقیه ژنوم). ۲۱.۸٪ از DMPs برای C3 در جزایر CpG در مقابل ۶.۹٪ و ۷.۴٪ برای C1 و C4، به ترتیب قرار داشتند. برای شناسایی قوی ترین و قابل توجه ترین علامت ژن های هیپو و هیپر متیله، برش ΔBeta را در ۰.۱۵ ثابت کردیم. با توجه به CpG های هیپومتیله، 13 DMP در C1، ۱۷ در C3 و ۱۱۹۴ در C4 یافت شد که به ترتیب مربوط به ۱۰، ۱۱ و ۷۶۱ ژن هیپومتیله است. پنج ژن با DMP های هیپومتیله در این ۳ خوشه مشترک بودند (PLSCR1، IFI44L، IFIT1، MX1، PARP9) (شکل ۵c)، که مربوط به ژن هایی است که برهمکنش های قوی ارائه می دهند (شکل ۵d). این ژن ها در مقایسه با HV به طور قابل توجهی در C2 هیپومتیله شدند. قابل توجه، ۵ ژن اضافی (HLA-A، DDX60، CMPK2، IFITM1)

۱۷ بیمار تحت درمان مقایسه کردیم. هیچ CpG با مقدار p تنظیم شده بنجامینی-هوشبرگ $FDR < 0.1$ به طور متفاوت در بیماران درمان شده در مقایسه با بیماران درمان نشده متیله نشد. یک هیپومتیلاسیون جهانی CpG برای همه خوشه ها مشاهده شد (۸۹.۶٪ در C1، ۱۰۰٪ در C2، ۶۷.۷٪ در C3 و ۸۰.۴٪ در C4). از آنجا که متیلاسیون DNA از نظر عملکردی مهم در نواحی پروموتور و در جزایر CpG20 رخ می دهد، توزیع DMP در مناطق مختلف ژنومی مورد بررسی قرار گرفت (شکل ۵b). نمایش بالاتری از DMPs در ناحیه پروموتور در C3 (۳۶.۴٪) و C1 (۳۳.۱٪) در مقایسه با C4 (۲۹.۱٪) یافت شد. نتیجه نمایش کمتر DMPs در مناطق بین ژنی برای C3 (۸.۸٪) در مقایسه با C1 (۲۲.۸٪) و C4 (۲۳.۱٪) بود. برای به دست آوردن بینش در مورد این الگو، ما پروب ها را بر اساس جزایر CpG تقسیم کردیم. سواحل (مناطق



شکل ۵: تجزیه و تحلیل متیلاسیون امضای قوی IFN در C1 و C3 را تأیید می کند و وضعیت متیلاسیون نابجا را در C4 نشان می دهد. آنالیز متیلاسیون خون کامل برای ۲۲۶ بیمار (C4: 26 و C3: 62 و C2: 57 و C1: 81) و pSS (C1: 81 و C2: 57 و C3: 62 و C4: 26) انجام شد که مقایسه های زوجی بین هر خوشه و HV انجام دادند. یک نمودار ون که همپوشانی سایت ها و ژن های متیله شده متفاوت را بین ۴ خوشه با ΔBeta مطلق > 0.075 نشان می دهد. توزیع DMP در مناطق مختلف ژنومی (بدنه ژنی، ۳' UTR، بین ژنی، ۵' UTR، IGR)، TSS، ۱۵۰۰ و ۲۰۰ TSS، و با توجه به تراکم CpG به جزیره CpG، قفسه، ساحل و باز دریا نمودار ون که همپوشانی CpG هیپومتیله و ژن ها را با ΔBeta مطلق > 0.15 بین سه خوشه IFN نشان می دهد. امضای IFN شبکه تعامل ۱۰ ژن مشترک C1 و C3 توسط تجزیه و تحلیل STRING با قطع اطمینان ۰.۴ هیپومتیله شده است. f . $FC \geq 1.5$ ، $p < 0.05$).

یک لنفوپنی که عمدتاً بر سلول های T تأثیر می گذارد در C1 یافت شد. در نهایت، متمایزترین خوشه از نظر توزیع و تعداد مطلق سلول ها C4 است. به طور خاص، C4 با درصدهای بالاتر و اعداد مطلق PMN (به ویژه نوتروفیل ها) در خون محیطی در مقایسه با سایر خوشه ها و HV مشخص شد. برعکس، درصد لنفوسیت ها (سلول های B و T) و مونوسیت ها در C4 در مقایسه با گروه شاهد یا سایر خوشه ها به طور قابل توجهی کاهش یافت. در نهایت، فرکانس های پایین تر و اعداد مطلق بازوفیل ها و DC ها نیز در این خوشه یافت شد.

سپس یک آنالیز دقیق از زیرجمعیت های سلولی مختلف انجام شد. اول، مونوسیت ها یک جمعیت سلولی ناهمگن را از نظر فنوتیپ و عملکرد نشان می دهند. بر اساس بیان CD14 و CD16، 3 زیر مجموعه مونوسیتی را می توان تعریف کرد که شامل کلاسیک (CD14++CD16-)، متوسط (CD14++CD16+) و غیر کلاسیک (CD14+CD16++) می شود. مونوسیت های کلاسیک برای پاسخ التهابی اولیه حیاتی هستند، می توانند به ماکروفاژها در بافت تمایز پیدا کنند و به بیماری مزمن کمک کنند. مونوسیت های میانی سلول های فاگوسیتی بسیار بالایی هستند که سطوح بالایی از ROS و واسطه های التهابی تولید می کنند. مونوسیت های غیر کلاسیک به طور گسترده ای به عنوان ضد التهاب تلقی می شوند، زیرا هموستاز عروقی را حفظ می کنند و اولین خط دفاعی را در شناسایی و پاکسازی پاتوژن ها تشکیل می دهند. جالب توجه است، فراوانی و تعداد مطلق مونوسیت های میانی در C1 و C3 افزایش یافت در حالی که فراوانی مونوسیت های کلاسیک در مقایسه با ۲ مورد دیگر کاهش یافت و زیرمجموعه غیر کلاسیک به طور قابل توجهی در C4 در راستای پاسخ التهابی مشاهده شده در این خوشه های مختلف کاهش یافت.

دوم، سلول های NK با بیان CD56 و فقدان کمپلکس CD3-TCR تعریف می شوند. علاوه بر این، بر اساس سطوح بیان CD16 و CD56، آنها در دو زیر مجموعه طبقه بندی می شوند: CD56hiCD16lo و CD56loCD16hi. زیرمجموعه سلول NK اخیر، سمیت سلولی طبیعی و وابسته به آنتی بادی را واسطه می کند و سطوح بالایی از پرفورین و افزایش کشتار را نشان می دهد. در مقابل، سلول های NK CD56hiCD16lo با سطوح پایین

و NLRC5 برای C1 و C3 مشترک بودند و همچنین به شدت با ژن های قبلی مرتبط بودند، و امضای IFN را در این دو خوشه تقویت می کردند (شکل ۵e). این ۱۰ ژن رایج هیپومتیل به پاسخ های دفاعی به ویروس نقش دارند و توسط IFN21 القا می شوند. ۷۵۶ ژن هیپومتیل به باقی مانده در C4 عمدتاً با مسیر دگرانولاسیون نوتروفیل مرتبط بودند. با توجه به CpG هایپر متیل، 41 DMP مربوط به ۲۵ ژن تنها در C4 یافت شد. این ژن ها عمدتاً در انتقال ZAP-70 به سیناپس ایمونولوژیک، فسفوریلاسیون زنجیره های CD3 از جمله زتا، فعال سازی پلاکت ها، سیگنال دهی و تجمع، هموستاز و سیگنال دهی PD-1 نقش دارند.

با ترکیب آنالیزهای رونویسی ($FC \geq 1.5$) و متیلومیک ($\Delta\text{Beta} \text{ مطلق} > 0.15$)، ترنسکرپتوم ۸، ۸ و ۱۲۶ ژن به ترتیب در ارتباط با کاهش وضعیت متیلاسیون در C1، C3 و C4 افزایش یافتند. ۵ ژن مشترک هیپومتیل شده قبلی که در سیگنال دهی IFN دخیل بودند نیز در سطوح رونویسی در ۳ خوشه بیان شده بودند. بیان بیش از حد رونوشت به شدت با هیپومتیلاسیون در C1 (8/10) و C3 (8/11) و به میزان کمتر در C4 (126/761) همراه بود. از بین ۱۲۶ ژن از C4، ۲۱ ژن در دگرانولاسیون نوتروفیل نقش داشتند که مطابق با ۲۲ Reactome Pathway Database (شکل ۵f) مرتبط ترین مسیرها را تشکیل می دهد. تنها ۶/۲۵ رونوشت در ارتباط با افزایش وضعیت متیلاسیون ژن های آنها در این خوشه سرکوب شد (CD247، CD3G، CDC25B، CXCR6، TBC1D4، UBASH3A).

آنالیز فلوسایتومتري

همانطور که تغییرات قابل توجهی در الگوهای لکوسیت های خون محیطی قبلاً شرح داده شده است، ما سپس ترکیب زیر مجموعه های لکوسیتی را در خوشه های مختلف بررسی کردیم. (شکل 6a&b). در C2، فرکانس و اعداد مطلق مشابه HV در تمام زیر مجموعه های مختلف تحلیل شده بودند. افزایش در فرکانس مونوسیت ها و لنفوسیت ها C3 را مشخص می کند، در ارتباط با افزایش قابل توجه فرکانس سلول های B. در همان زمان،



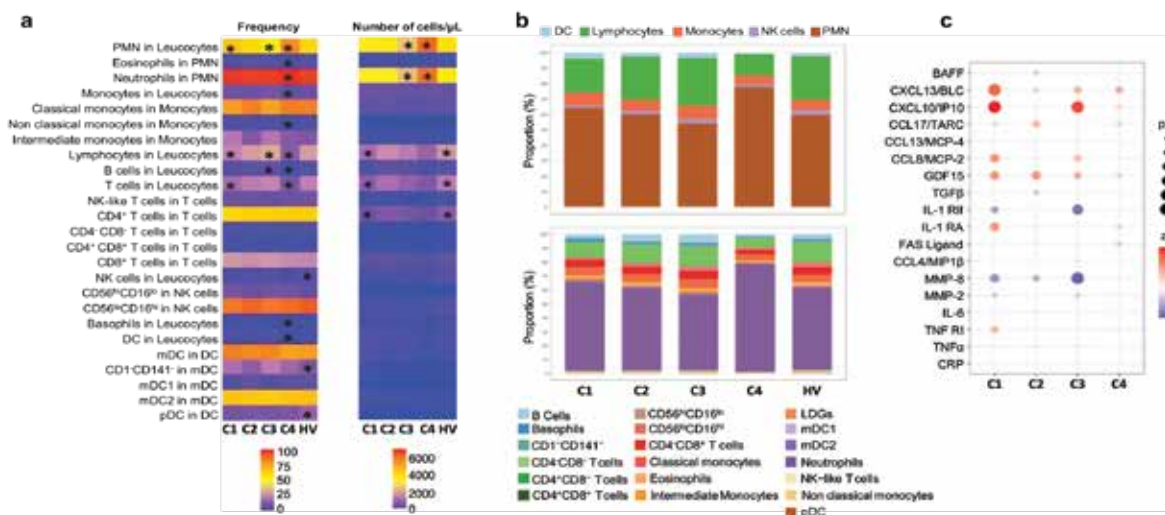
IFN قوی، افزایش یافتند. در همان زمان، IL-1 RII، گیرنده طعمه برای سیتوکین متعلق به خانواده IL-1، در C1 و C3 تنظیم شد. به طور کلی، C1 تا حد زیادی در IL-6، CXCL13/BLC، IL-1RA و IL-1 غنی شد. سطوح MMP-8، یک پروتئاز که عمدتاً توسط نوتروفیل ها بیان می شود، متفاوت از HV در C4 نبود اما در خوشه های دیگر کمتر بود. توجه داشته باشید، بسیاری از سایتوکاین ها مانند CXCL10/ IP-10، CXCL13/BLC، BAFF همه خوشه ها از جمله C2 در مقایسه با HV افزایش یافته اند. با این حال، هیچ تفاوتی بین خوشه ها برای CRP، Fas Ligand، CCL13/MCP-4، CCL4/MIP-1 β و TGF β یافت نشد.

برای تأیید اینکه بیماران دارای امضای فعال IFN برای نوع I در گردش هستند، ما سطوح IFN α در پلاسما را با استفاده از فناوری آرایه تک مولکولی Simoa در بیماران pSS و HV اندازه گیری کردیم. سطح متوسط IFN α در پلاسما ۸۰۷ (۱۷۴-۱۷۷) fg/ml و ۵۳۰ (۱۰۳-۱۰۶) fg/ml در C1 و C3 بود، در حالی که

پرفورین مشخص می شوند و در درجه اول برای تولید سیتوکین از جمله IFN26،27 تخصصی هستند. بر این اساس، فرکانس سلول های NK CD56hiCD16lo از CD56loCD16hi در C1، C3 و C4 و به میزان کمتری در C2 افزایش یافت. این ممکن است تا حدی تنظیم بالا سیتوکین ها و مسیرهای اینترفرون را در خوشه های بیماری توضیح دهد. اگرچه تصور می شود که سلول های دندریتیک پلاسما سیتوئید (pDCs) نمایانگر سلول های اصلی تولیدکننده IFN α هستند، هیچ تفاوتی بین خوشه ها مشاهده نشد و کاهش آنها در خون محیطی بیماران pSS در مقایسه با HV28 تأیید شد.

آنالیز سیتوکین

ما متعاقباً ارزیابی کردیم که آیا خوشه های pSS نیز تفاوت هایی را در پارامترهای سیستمیک التهاب، مانند سیتوکین ها، کموکاین ها و سایر عوامل محلول نشان می دهند (شکل ۶c). پروتئین القا شده با IFN γ (CXCL10/IP-10) و همچنین CCL8/MCP-2 و TNF α در C1 و C3، یعنی دو خوشه اصلی مرتبط با امضای



شکل ۶: توزیع زیر مجموعه سلولی در خون و سیتوکین ها، کموکاین ها و واسطه های التهابی در سرم در ۴ خوشه و داوطلبان سالم (HV). آنالیز فلوسایتومتری برای بیمار C1: 36) و C4: 80؛ C2: 71؛ C3: 96 و HV ۳۰۹ انجام شد. ۲ نقشه حرارتی توزیع میانگین زیرمجموعه های سلول های خونی را در فرکانس (%۱۰۰-۰) و در اعداد مطلق (در هر میکرولیتر خون) در ۴ خوشه و HV ارزیابی شده توسط فلوسایتومتری نشان می دهد. ستون ها خوشه ها و HV را نشان می دهند و ردیف ها زیر مجموعه های مختلف سلول هستند. ستاره به این معنی است که خوشه (یا HV) از نظر آماری با بقیه متفاوت است. b داده های فلوسایتومتری نشان داده شده توسط نمودارهای میله ای نسبت انواع سلول در هر خوشه. c واسطه های سرمی برای ۱۹۲ بیمار (C1: 67؛ C2: 48؛ C3: 61؛ C4: 16) و pSS (HV ۱۷۱) مورد تجزیه و تحلیل قرار گرفت. توزیع بیمار و HV با توجه به هر متغیر تجزیه و تحلیل شده در روش ها توضیح داده شده است. TNF-RII، IL-1 RII، TNF-RI، CXCL13/BLC، FAS Ligand، GDF-15، CXCL10/IP-10، CCL8/MCP-2، CCL13/MCP-4، CCL4/MIP-1 β ، MMP-8، CCL17/TARC، IL-6، BAFF بیان شدند. محلول BAFF، IL-6، TNF α ، CRP، MMP-2، و TGF β با روش ایمونواسی آنزیم ساندویچ کمی اندازه گیری شد و به صورت pg/ml بیان شد. سطوح غلظت سیتوکین یا کموکاین برای هر خوشه با HV مقایسه شد. اهمیت آماری با استفاده از یک آنالیز واریانس یک طرفه و به دنبال آن آزمون توکی پس از آن تعیین می شود. اهمیت بین خوشه و HV به صورت گلوله ای از کوچک (غیر معنی دار) تا بزرگ (معنی دار) نشان داده می شود. جهت ارتباط به عنوان z-score نشان داده می شود که در آن گلوله قرمز به سمت بالا تنظیم می شود و گلوله آبی پایین تنظیم می شود.

($p\text{-value} = 0.016$). بین چهار خوشه مشاهده شد. همچنین بیماران C4 علائم بالینی شدیدتری را در مقایسه با سه خوشه دیگر گزارش کردند. برخی از ویژگی‌های سرولوژیکی در ۴ خوشه تفاوت معنی داری داشتند، هیپرگاماگلوبولینمی ($p\text{-value} < 0.001$) (جدول ۲)، آنتی بادی‌های آنتی ژن هسته ای قابل استخراج ($p\text{-value} < 0.001$) (ENA)، وجود سرمی anti-SSA52/anti-SSA60 اتوانتی بادی ها ($p\text{-value} < 0.001$) و زنجیره‌های سبک آزاد کاپا و لامبدا با گردش بالاتر ($p\text{-value} < 0.001$) (cFLC) (شکل C1). C3 با سطوح بالاتری از این پارامترها در مقایسه با C2 و C4 همراه بودند. علاوه بر این، C2 و C4 در بیماران مبتلا به تظاهرات غددی بیماری که در غیاب آنتی‌بادی‌های ضد SSA با امتیاز تمرکز مثبت ارزیابی شده‌اند، غنی شدند (جدول ۲). علاوه بر این، سطوح فاکتور روماتوئید ($p\text{-value} < 0.001$) و سطوح کسر مکمل ($p\text{-value} = 0.003$) C4 از نظر آماری بین چهار خوشه متفاوت بود. C1 با فاکتور روماتوئید بالاتر و کاهش سطح کسر C4 مکمل در مقایسه با خوشه‌های دیگر مشخص شد. در حالی که برخی از بیماران آنتی بادی‌های ضد dsDNA را در C1 و C3 و آنتی بادی‌های ضد CCP را در C4 ارائه کردند، تقریباً هیچ یک از این آنتی بادی‌های اتوانتی بادی در سایر خوشه ها وجود نداشت.

پیش‌بینی عضویت بیمار در هر یک از چهار خوشه

سپس از طریق رویکردهای یادگیری ماشینی، یک مدل ترکیبی ایجاد کردیم که می‌تواند با توجه به تعداد کمی از متغیرها، پیش‌بینی کند که هر بیمار به کدام یک از ۴ خوشه تعلق دارد. مدل ترکیبی پیشنهادی با یک رویکرد ۲ مرحله‌ای برای تخصیص بیماران به خوشه درست ساخته شد. مجموعه نهایی ویژگی‌های انتخاب شده از ۱۰ ژن برای مدل پیش‌بینی C4 (مرحله اول) و ۳۱ ژن برای مدل طبقه بندی C1، C2 و C3 (مرحله دوم) تشکیل شده است. توزیع بین خوشه‌های بیان نرمال شده تبدیل تثبیت کننده واریانس (vst) برای همه این رونوشت‌ها در شکل تکمیلی ۱۱ نشان داده شده است. مجموعه اعتبارسنجی (شکل ۱ و جدول ۱) برای آموزش

سطوح در گردش در سایر خوشه‌ها و HV نزدیک به حد پایین اندازه‌گیری بود. جالب توجه است که IFN α در سرم با دو ماژول ترنسکریپتوم M۱.۲ و IFN و ماژول IFN α که در شکل ۳ توضیح داده شده است، به ویژه در C1 و به میزان کمتر در C3 همبستگی مثبت داشت، که تایید کننده امضای IFN نوع I مشاهده شده در این بیماران است. نکته قابل توجه، نیمی از بیماران در C2 ضد مالاریا دریافت کردند و مطالعات قبلی نیز نشان داده اند که استفاده از هیدروکسی کلروکین می‌تواند سطح گردش خون نوع I29,30 و نوع II31,32 را کاهش دهد. IFN z-scores. IFN α در سرم با ESSDAI ارتباط نداشت. اما سطوح بالاتر IFN α سرم با حوزه‌های خونی و بیولوژیکی ESSDAI مرتبط بود.

علائم بالینی و ویژگی‌های سرولوژیکی

سابقه پزشکی بیمار و ویژگی‌های بیماری شامل پارامترهای بالینی و سرولوژیکی برای ۳۰۴ بیمار pSS جمع آوری شد. بیماران از C2 در مقایسه با بیماران از سایر خوشه ها، مدت زمان بیماری کمتری داشتند. اگرچه ارزیابی جهانی پزشک (PGA) برای کل جمعیت جمع‌آوری شد، ESSDAI و ESSPRI فقط در مراکز خبره (بارسلونا، برست، کوردوبا، ژنو، هانوفر، لوون، میلانو، پورتو و سگد) در زیر مجموعه‌ای از ۱۹۳ و ۱۵۰ مورد ارزیابی قرار گرفتند.

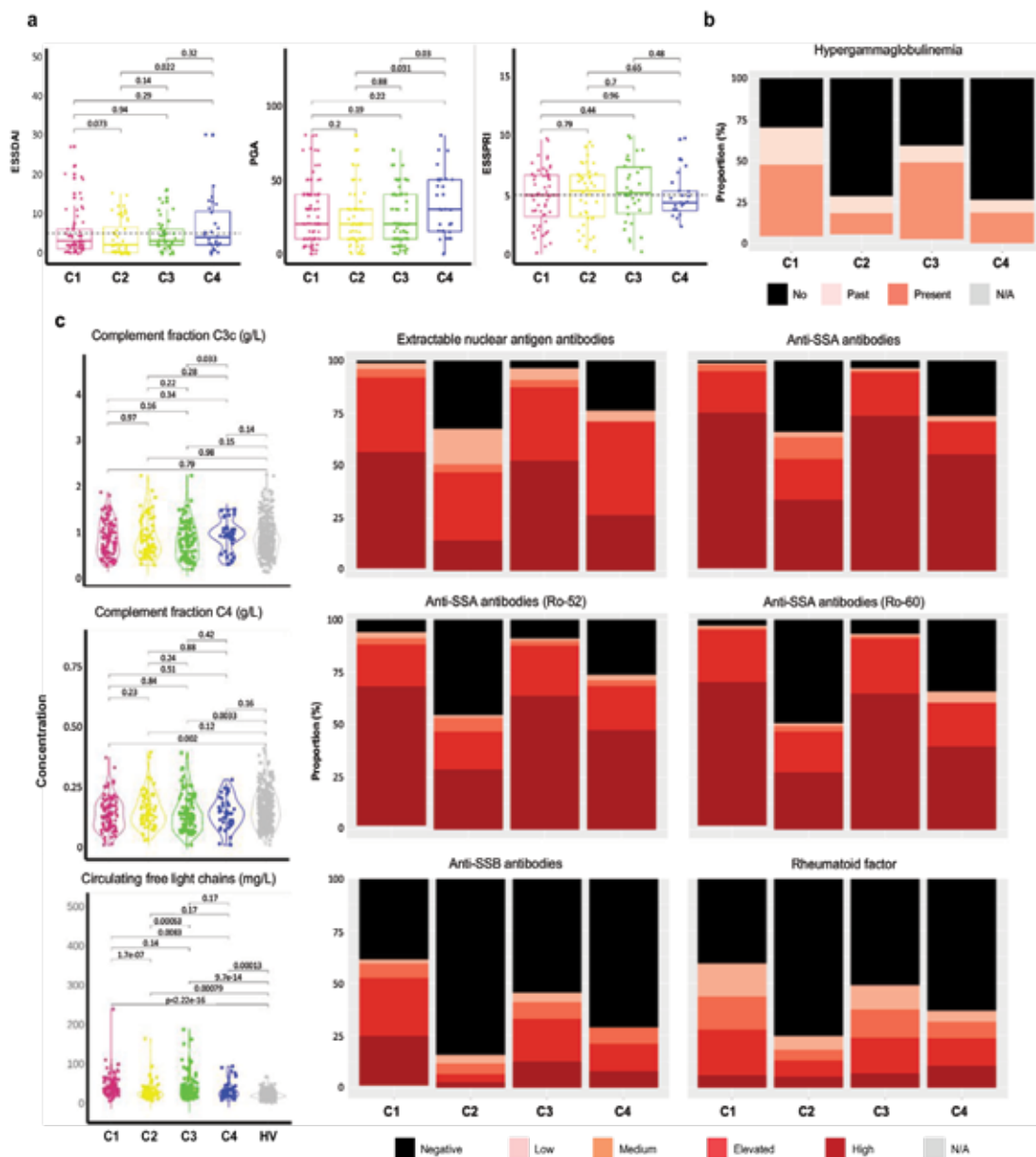
کمترین میانگین نمره ESSDAI در C2 و بالاترین میانگین نمرات ESSDAI و PGA در C4 مشاهده شد (جدول ۲، شکل 7a) اما از نظر آماری تفاوت معنی داری بین ۴ خوشه وجود نداشت. هیچ تفاوت واضحی در اجزای ESSDAI و یا در معیارهای عینی خشکی چشم و بزاق بین ۴ خوشه مشاهده نشد. علاوه بر این، هیچ تفاوت معنی داری برای نمره جهانی ESSPRI و ۳ مولفه آن (یعنی خشکی، درد و خشکی) وجود نداشت، به جز بین بیماران C2 مثبت SSA که نمرات ESSPRI پایین تری ($p\text{-value} = 0.001$) در مقایسه با بیماران SSA منفی گزارش کردند.

تفاوت آماری معنی داری در توزیع آرتریت گزارش شده ($p\text{-value} = 0.016$)، میزان سابقه سرطان ($p\text{-value} = 0.028$)، بیماری عروق کرونر ($p\text{-value} = 0.002$) و بیماری انسدادی مزمن ریه



تمایز مربوطه، و احتمال تعلق به یکی از ۴ خوشه به ترتیب در شکل ۸a و b، نشان داده شده است. ما از یک گروه اولیه مستقل از ۳۷ بیمار pSS برای تعمیم مدل ترکیبی استفاده کردیم. پس از پیش بینی، C1

به دلیل ناهمگونی بیماران pSS C4 در این مجموعه، و مدل ترکیبی سپس بر روی مجموعه کشف اجرا شد. دقت مدل برای مرحله اول و دوم به ترتیب ۹۵.۱۵٪ با ۹۹.۱۲٪ و ۹۵.۵۷٪ بود. ماتریس سردرگمی، تحلیل تابع



شکل ۷: فعالیت بیماری و توزیع سرولوژیکی در ۴ خوشه. a) جمع آوری شده برای ۱۹۳ بیمار PGA (C1: 70, C2: 52, C3: 44, C4: 27) و PSS جمع آوری شده برای ۲۸۶ بیمار (C1: 94, C2: 71, C3: 85, C4: 36) و توزیع شده برای ۱۵۰ بیمار PSS (C1: 56, C2: 43, C3: 30, C4: 21) توزیع در ۴ خوشه نشان داده شده است. نتایج آزمون مجموعه رتبه ویلکاکسون جفتی Two-tailed نشان داده شده است. b) باریتات نسبت هیپرگامگلوبولینمی گذشته (نارنجی روشن) یا فعلی (نارنجی) (C1: 97, C2: 73, C3: 86, C4: 38) را در هر خوشه نشان می دهد. c) آنتی بادی های آنتی ژن هسته ای قابل استخراج، آنتی بادی های ضد SSA، آنتی بادی های ضد SSA (RO-۶۰)، آنتی بادی های ضد SSB، فاکتور روماتوئید برای ۳۰۴ بیمار PSS انجام شد (C1: 101, C2: 77, C3: 88, C4: 38) و HV ۲۳۰ و در سرم، در همان مرکز، با استفاده از یک ایمونوآنالیز شیمیایی خودکار (IDS-ISYS) اندازه گیری می شوند. باریتات ها نسبت سطح غلظت را در هر خوشه نشان می دهد (سیاه: منفی، صورتی روشن: کم، نارنجی: متوسط، قرمز: بلند و قرمز تیره: زیاد). از توریدیمتری برای فاکتور روماتوئید (RF)، کسری مکمل C3C و C4 تعیین شده و زنجیره های نور آزاد استفاده شد. اهمیت آماری با آزمون جمع دو برابر بلند و Wilcoxon-Rank دو دم دو دم تعیین می شود. توطئه ها با میله های خطا که نشان دهنده دامنه interquartile است، میانه را نشان می دهد. توزیع بیمار و HV با توجه به PGA و پارامترهای بیولوژیکی متغیر تجزیه و تحلیل در روش ها توضیح داده شده است.

تحلیل موجود در مخزن GitHub آزمایشگاه ۳۳ یکپارچه شده است.

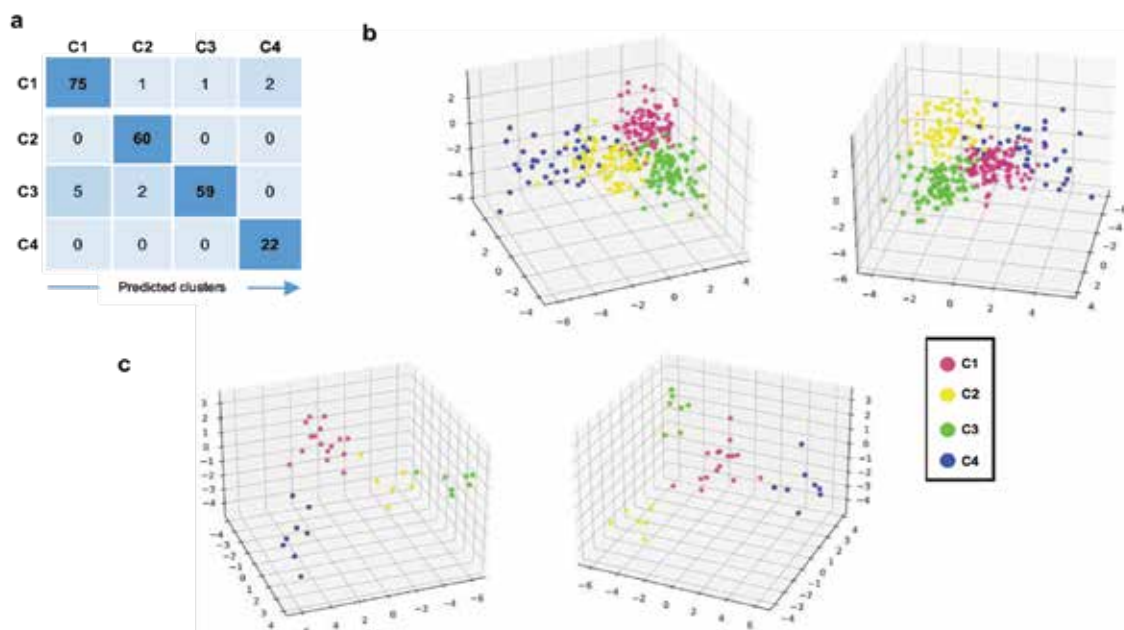
بحث

در طول دهه گذشته، بسیاری از درمان های تعدیل کننده ایمنی هدفمند برای pSS نتوانسته اند در کارآزمایی های بالینی سودمند باشند. از این رو هنوز هیچ درمان اصلاح کننده بیماری برای این بیماری تایید نشده است. ماهیت ناهمگن pSS و توسعه غیرخطی آن، همراه با شعله ور شدن فعالیت و بهبودی متعاقب آن همراه با تظاهرات بالینی بسیار ناهمگون، ممکن است شکست های کارآزمایی بالینی را توضیح دهد. در این زمینه، علاقه فزاینده ای به شناسایی زیرگروه های مشخص شده از بیماران، پیش نیاز عنوان بیومارکرهای مولکولی پیش بینی کننده پاسخ درمانی وجود دارد.

در اینجا در مورد یک مطالعه پروفایل مولکولی قابل توجهی که بر روی بیماران pSS انجام شده است را بدون توجه به فنوتیپ های بالینی گزارش می کنیم. مطالعات قبلی در pSS عمدتاً بر درگیری سیگنالینگ IFN تمرکز دارد. بنابراین، بیماران pSS را می توان به

شامل ۱۶ بیمار (۴۳.۲٪)، C2 ۶ بیمار (۱۶.۲٪)، C3 ۷ بیمار (۱۸.۹٪) و C4 ۸ بیمار (۲۱.۶٪) بود. آنالیز تابع متمایز مربوطه و احتمال تعلق یک بیمار به یکی از ۴ خوشه به ترتیب در شکل ۸c و داده های تکمیلی ۱۰ نشان داده شده است. سپس از فهرست حداقل ۲۵۷ امضای ژن متمایز که قبلاً در شکل ۱a انتخاب شده بود برای ایجاد یک نقشه حرارتی با پیش بینی ایجاد شده توسط مدل ترکیبی استفاده کردیم. خوشه های مشاهده شده دارای مشخصات مشابهی با آنهایی بودند که در مجموعه کشف شناسایی شده بودند و دوباره در مجموعه اعتبارسنجی مشاهده شدند (شکل ۱a)، که یک بار دیگر مدل خوشه بندی را تایید می کند. علاوه بر این، بیماران پیش بینی شده توزیعی از امضای IFN را نشان دادند مطابق با امضای خوشه های شناسایی شده (شکل ۳). در مجموع، این مشاهدات اعتبار مدل ترکیبی ما را تقویت می کند.

در نهایت، برای اینکه مدل ما بتواند گروه های دیگر بیماران را پردازش کند، یک تابع درون یابی مبتنی بر شش ژن را اجرا می کنیم که بیان ثابتی را در هر ۴ خوشه و HV ارائه می کند. مدل ترکیبی در یک ابزار تجزیه و



شکل ۸: توسعه یک مدل ترکیبی برای پیش بینی تعلق یک بیمار به یکی از ۴ خوشه. a. ماتریس سردرگمی از مدل ترکیبی در گروه کشف انجام شده برای ۲۲۷ بیمار (C1: بیمار ۲۲۷، C2: ۶۰، C3: ۶۶ و C4: ۲۲) نشان داده شده است. b. آنالیز عملکرد متمایز (DFA) بیماران پیش بینی شده از گروه کشف، خوشه های کاملاً جدا شده را نشان می دهد. دو نمای مختلف از یک DFA نشان داده شده است. c. DFA بیماران پیش بینی شده از گروه ابتدایی، خوشه های کاملاً جدا شده را نشان می دهد. دو نمای مختلف از یک DFA نشان داده شده است. سی و هفت بیمار pSS از گروه ابتدایی مورد آنالیز قرار گرفتند و به صورت C4: ۸ و C1: ۱۶، C2: ۶، C3: ۷ و C4: ۸ پیش بینی شدند.



به غده بزاقی در آینده تصویر کاملی از پاتوفیزیولوژی بیماری به خصوص در C2 ارائه می دهد.

سه خوشه دیگر تفاوت های قابل توجهی با HV و به ویژه امضای ژن برجسته IFN نشان دادند. این یافته ها به شواهد روبه رشد نقش مهم مسیرهای IFN در پاتوژنز اختلالات سیستمیک و اندام خاص از جمله pSS اضافه می کنند. در حالی که IFN نوع I به عنوان عوامل اصلی در پاتوژنز pSS پیشنهاد شده بود، نقش IFN نوع II در پاتوژنز بیماری نیز برجسته شده است ۴۷، ۶. نتایج ما نشان می دهد که امضای IFN در ۳ خوشه مبتنی بر IFN متفاوت است. بیماران C1 بالاترین نمرات IFN نوع I و II را داشتند، C3 امتیاز IFN نوع I بالاتر از C4 داشتند و این ۲ خوشه نمرات IFN نوع II مشابهی داشتند. بنابراین، امتیاز IFN C4 عمدتاً توسط فعال سازی IFN نوع II هدایت می شود. در نتیجه، C1 و C3 با خوشه IFN توصیف شده توسط جیمز و همکاران یکسان بودند، همچنین با سطوح بالای پروتئین خون CXCL10/IP-10 مرتبط بود.

مطابق با نمرات مشاهده شده IFN، سطح سرمی IFN α در گردش با امضای IFN نوع I (شکل ۳)، به ویژه در C1 و به میزان کمتر در C3، همبستگی مثبت داشت. با این حال، سطوح IFN α در سرم با نمره جهانی ESSDAI همبستگی نداشت، اما سطوح بالاتر IFN α سرم با حوزه های خونی و بیولوژیکی ESSDAI مرتبط بود.

در حالی که IFN عمدتاً C1 را هدایت می کند، افزایش فراوانی لنفوسیت های B در خون مرتبط با فعال سازی قابل توجه مسیرهای متعارف مرتبط با فعال سازی سلول B، مانند سیگنال دهی گیرنده سلول B و توسعه سلول B در C3 مشاهده شد. ویژگی های اصلی بیولوژیکی مرتبط با C3 و همچنین C1 هیپرگاماگلوبولینمی، آنتی بادی های ضد هسته ای، آنتی بادی های ضد SSA52/anti-SSA60، سرم و cFLC بالاتر بود که آنچه قبلاً در بیماران pSS مثبت آنتی بادی گزارش شده بود را تأیید می کند. در نهایت، SNP های مرتبط با ژن های HLA کلاس II عمدتاً در بیماران C1 و C3 گزارش شده اند که دارای امضای مثبت IFN و سطوح بالای اتوآنتی بادی هستند، همانطور که در SLE۱۶ نشان داده شده است.

بیماران از C4 فنوتیپ بالینی شدیدتری را نسبت به سایرین نشان دادند، با امضای رونویسی التهابی که عمدتاً

زیرگروه های اینترفرون منفی، نوع I یا نوع I+II مثبت با شیوع بالاتر ضد SSA و anti-SSB در میان افرادی که فعال سازی IFN بدون ارتباط با فعالیت سیستمیک دارند، طبقه بندی کرد. گروه دیگری آنالیز خوشه بندی ریزآرایه بیان ژن خون را انجام دادند، که ۴۷ بیمار pSS را در سه خوشه طبقه بندی کرد که با IFN و التهاب بدون ویژگی های بالینی متمایز مشخص می شد. علاوه بر این، چهار زیر گروه از بیماران با ویژگی های بالینی مشابه بر اساس تعداد سلول های مطلق در هر میکرولیتر خون شناسایی شدند. در نهایت، یک طبقه بندی بر اساس فنوتیپ های بالینی بیمار که به صورت پسینی در سطح مولکولی مشخص می شود، پیشنهاد شد. این کارها مبنای معقولی برای ساختن یک طبقه بندی مولکولی pSS فراهم می کند. رویکرد یکپارچه ما با استفاده از ویژگی های بالینی مولتی امیکس و بیمار امکان درک بیشتر ناهمگونی pSS را فراهم می کند.

ما ماژول های ترنسکریپتوم را شناسایی کردیم که به ما امکان می دهد بیماران pSS را بدون توجه به نوع درمانشان، به چهار دسته مجزا جدا کنیم، که الگوهای خاصی از اختلالات سیستم ایمنی را با فعالیت بیماری و میانگین علائم گزارش شده توسط بیمار، نمرات مشابه با گروه های طبیعی مانند ASSESS و UKPSSR منعکس می کند.

بیماران از C2 نمایه ای مانند سالم نشان دادند که با این وجود شامل بیماران pSS است که سطح مشابهی از علائم عینی خشکی، درد و خستگی را گزارش می کنند (البته ESSDAI پایین تری نسبت به سه خوشه دیگر). C2 همچنین در بیماران مبتلا به تظاهرات بیماری غدد با امتیاز تمرکز مثبت و بدون آنتی بادی ضد SSA غنی شد. خوشه مشابهی اخیراً بدون افزایش در ماژول های IFN و حداقل فعالیت ماژول های ژنی مرتبط با التهاب توصیف شد. تمام داده های پروفایل مولکولی گزارش شده در اینجا از نمونه های خون به دست آمده است که می تواند بر تفسیر برخی از نتایج تأثیر بگذارد. به عنوان مثال، کاهش pDC های خون محیطی بیماران pSS در مقایسه با HV که قبلاً گزارش شده است، به این معنا نیست که pDC در غدد بزاقی غنی شده است و این احتمال وجود دارد که محل های بافتی ممکن است منبع مهم IFN α در این افراد باشد. گسترش آن آنالیزها

از آنجایی که فرضیه ها از یک مطالعه مقطعی و یک گروه ابتدایی کوچک استخراج شده اند، یافته ها باید در گروه های طولی تأیید شوند تا مشخص شود که آیا بیماران به صورت طولی در خوشه اولیه خود باقی خواهند ماند؟ (صرف نظر از سطح فعالیت بیماری و درمان های دریافتی، یا اینکه آیا درمان ها کاهش فعالیت بیماری با اصلاح وسعت و دامنه اختلالات سیگنال دهی ژن). در مجموع، نتایج ما می تواند استراتژی های درمان pSS را بهبود بخشد و به یک رویکرد بیمار محور اجازه دهد. این پارادایم که قبلاً در زمینه انکولوژی اجرا شده است، احتمال موفقیت آزمایشی را افزایش می دهد و توسعه داروهای کارآمد جدید علیه pSS را تقویت می کند.

منبع:

<https://www.nature.com/articles/s41467-021-23472-7>

به سیگنال دهی سیتوکین از پاسخ فاز حاد مرتبط بود. C4 همچنین با لنفونی عظیم و سطوح بالای نوتروفیل ها مشخص شد. قبلاً نشان داده شده بود که نسبت نوتروفیل به لنفوسیت (NLR) با فعالیت بیماری در خودایمنی سیستمیک ارتباط دارد. در واقع، در مطالعه ای بر روی ۴۸۳ بیمار بالغ مبتلا به مولتیپل اسکلروزیس، NLR می تواند بین مولتیپل اسکلروزیس عودکننده - فروکش کننده و مولتیپل اسکلروز اولیه پیشرونده تمایز قائل شود و بدتر شدن ناتوانی را پیش بینی کند. مطالعات بیشتری در pSS برای ارزیابی اهمیت این نسبت مورد نیاز است. از آنجایی که چالش اصلی فعلی در کارآزمایی های بالینی درمان های جدید برای pSS، انتخاب بیماران مناسب است، ما در اینجا ترکیبی از پارامترهای مولکولی را پیشنهاد می کنیم که امکان طبقه بندی بیمار بر اساس اندوتیپ ها را فراهم می کند. سپس ما یک مدل ترکیبی برگرفته از یادگیری ماشینی را بر اساس استفاده از تعداد محدودی رونوشت از RNASeq خون کامل و اعتبارسنجی در مجموعه داده های مستقل از یک مطالعه اولیه pSS ایجاد کردیم تا امکان تحلیل مجدد کارآزمایی های بالینی قبلی و در حال انجام را فراهم کنیم. پیش بینی کننده های پاسخ درمانی را به تصویر می کشد.

این یافته ها پیامدهای قابل توجهی برای درمان بیماران pSS دارند، و منطقی برای موقعیت یابی بهینه دارو و ترکیبی از داروها با مکانیسم های اثر مکمل ارائه می کنند. به طور خاص، یافته های ما یک منطق قوی برای درمان بیماران دارای پروفایل C1، C3، یا C4 با مهارکننده های پاسخ های IFN نوع I به تنهایی یا به صورت ترکیبی ارائه می کند، زیرا از ارتباط سلول های B به عنوان اهداف بالقوه درمانی در بیماران C3 پشتیبانی می کنند. کارآزمایی ها با آنتی بادی های کاهش دهنده سلول B (rituximab) نتایج امیدوارکننده ای را در کاهش فعالیت سیستمیک در pSS51 نشان داده اند.

مناطق که نیاز به بررسی بیشتر دارند شناسایی شده است. اولاً، اگرچه امضاهای ژن خوشه ای شناسایی شده ما برای غلبه بر عدم تعادل در تعداد سلول های خونی به اندازه کافی قوی هستند و با طول مدت بیماری مرتبط نیستند، به جز C2، تجزیه و تحلیل RNA-Seq از ترکیب نوع سلولی نمونه غافل است. تحقیقات بیشتر با استفاده از رویکردهای deconvolution ادامه دارد. دوم،