مدل گرافیکی احتمالی برای کشف جامع امضاهای جهش‌زا

مینا شایگان, سید علی کتان‌فروش

۴تیر, ۱۳۹۹

چکیده

شناخت فرایندهای جهش‌زایی که در طی گسترش سرطان عمل می‌کنند، موضوع اصلی بیولوژی سرطان است. جهش‌های جسمی ناشی از این فرایندها با تکثیر دی‌ان‌ای یا قرار گرفتن در معرض عوامل محیطی و روش زندگی مرتبط است. دانستن فعالیت فرایندهای جهش‌زا در یک تومور می‌تواند در روش‌های درمانی شخصی‌سازی شده، تشخیص زودهنگام، و تحقیقات تومورشناسی مورد استفاده قرارگیرد. روش‌های محاسباتی ۳۰ امضای معتبر از فرایندهای جهش‌زا فعال در سرطان‌های انسانی شناسایی کرده‌اند. هر امضا یک الگو از انواع جهش است. هرچند که مشخص است که امضاهای مختلف از بافت‌های مختلفی بدست آمده‌اند که ممکن است ضایعه اولیه یکسان داشته باشند. برای تعیین یک امضای جهش‌زا متناسب با یک فرایند جهش‌زا لازم است امضاهای بدست آمده از همه ضایعات اولیه‌ای که فرایند جهش‌زا در آن عمل کرده است را با یکدیگر ادغام کنیم. نمونه‌های هر ضایعه اولیه نباید جداگانه بررسی کرد بلکه باید همه نمونه‌ها با هم در نظر بگیریم. اما این کار پیچیدگی مجموعه داده را افزایش داده و شناسایی امضاء را دشوار می‌کند. برای حل این مشکل از یک مدل سلسله‌مراتبی تخصیص دیریکله پنهان همراه با استنباط بیزی وردشی جهت استخراج امضاهای جهش‌زا استفاده می‌شود تا بتوانیم ویژگی‌های هر ضایعه اولیه بدست آوریم. از کران‌های پایین وردشی جهت پیدا کردن تعداد الگوهای جهش ممکن استفاده می‌شود.

کلمات کلیدی: فرایند جهش‌زا، امضا جهش، سرطان، مدل گرافیکی احتمال

مقدمه

فرایندهای جهش‌زا متعددی از جمله خطاهای تکثیر دی‌ان‌ای و قرار گرفتن در معرض عوامل جهش‌زایی مانند مواد شیمیایی، رادیواکتیو، دخانیات سرطان زا [۱]، نور فرابنفش [۲]، و واکنشهای التهابی باعث ایجاد جهش‌ها هستند. جهش‌های محرک[۱] که در فرایند تبدیل شدن سلول سالم به سرطانی دخیل هستند، قابلیت رشد را به سلول‌های سرطانی می‌دهند و در ریزمحیط بافتی که سرطان ایجاد می‌شود به طور مثبت انتخاب می‌شوند (نسبت جهش‌هایی که یک آمینواسید را تغییر می‌دهند به نسبت جهش‌هایی که آمینواسیدی را تغییر نمی‌دهند, بیش‌تر از انتظار است). برای حفظ نهایی سرطان نیازی به یک جهش محرک نیست (اگرچه همچنان حضور دارند) اما باید در بعضی از مراحل پیشرفت سرطان انتخاب شوند. در مقابل جهش‌های گذرگر[۲] به طور مثبت انتخاب نمی‌شوند و قابلیت رشد تجمعی را ایجاد نمی‌کند و در نتیجه در پیشرفت سرطان دخالت ندارد. جهش‌های گذرگر در ژنوم سرطانی یافت می‌شود زیرا جهش‌های جسمی بدون عواقب عملکردی اغلب در طول تقسیم سلولی رخ می‌دهند. بنابراین یک سلولی که یک جهش محرک بدست می‌آورد، از قبل جهش‌های جسمی بی‌اثر زیستی در ژنوم خود دارد. این جهش‌ها به دنبال توسعه تجمعی همراه می‌شوند و در نتیجه در تمام سلول‌های سرطان نهایی حضور دارند [۳]. مطالعات ژنوم سرطانی معمولا بر روی شناسایی جهش‌های محرک تمرکز دارند. این جهش‌ها در درک مکانیسم توسعه سرطان کمک می‌کنند. با این حال جهش‌های گذرگر نیز اطلاعات مهمی در اختیار ما قرار می‌دهند، زیرا که اغلب آنها الگوهایی (امضاهای جهش‌زایی) را نشان می‌دهند که باعث ایجاد جهش‌های جسمی می‌شوند. مطالعات کلاسیک با توجه به محدودیت توالی‌یابی برای درک فرایندهای جهش جسمی بر روی ژن‌های سرطانی محدودی تمرکز می‌کردند که فراوانی جهش در آنها زیاد است. این جهش‌ها در افراد مختلف با یک نوع سرطان جهت به دست آوردن جهش‌های کافی برای تجزیه و تحلیل جمع‌آوری می‌شوند. سپس پروفایل‌های الگو جهش با انواع مختلف سرطان با یکدیگر مقایسه می‌شوند. با این حال، از آنجایی که

---

[۱] driver mutation

[۲] passenger mutation

بسیاری از جهش‌ها در ژن‌های سرطانی جهش‌های محرک ناشی از تکثیر سلولی هستند، پروفایل جهش حاصل یک نمایش سودار[۱] از فرایند جهش متناظر است. علاوه بر این کمبود داده‌های جهش امکان مقایسه الگوهای جهش افراد وجود ندارد. به دلیل توسعه توالی‌یابی نسل جدید[۲]، داده‌هایی با مقیاس بزرگ از ژنوم سرطانی به سرعت در دسترس همگان قرار گرفته است و فرصت‌های جدیدی را برای بررسی امضاهای جهش به صورت فردی با یک نمایش ناسودار با استفاده از داده‌های جهش جسمی گسترده ژنوم فراهم می‌کند[۴].

هر فرآیند جهش‌زا منجر به امضای جهش‌زا خاصی می‌شود و انباشت جهش‌ها از بدو تولد تا به امروز نتیجه ترکیبی از برخی از امضاهای جهش‌زا در نظر گرفته می‌شود [۵]. از این رو ، انتظار می‌رود که تفکیک‌سازی امضاهای جهش‌زا مکانیسم‌های سرطان‌زا را آشکار کند و ممکن است به عنوان نشانگرهای زیستی برای تشخیص زودهنگام عمل کنند [۶،۷،۸]. اخیرا با استفاده از آزمایش‌های مبتنی بر ویرایش ژن CRISPR-Cas9 در یک سیستم سلول مهندسی‌شده ژنتیکی[۳] انسان، مجدداً امضاهای جهش‌زا سرطانی در شرایط آزمایشگاهی[۴] بازتولید شدند. این نتیجه از وجود امضاهای جهش‌زا پشتیبانی می‌کند [۹].

امضاها با شناسایی الگوهای مشترک در بسیاری از تومورها براساس تعداد جهش‌ها و نوع آنها کشف می‌شوند. روش اصلی کشف امضا براساس فاکتورگیری نامنفی ماتریس[۵] بود [۱۰] و ۳۰ امضاء جهش‌زا در انواع مختلف سرطان آشکار کرد [۱۱،۱۲،۱۳،۱۴][۶]. رویکرد "signeR" نیز از روش فاکتورگیری نامنفی ماتریس در ترکیب با رویکرد بیزی تجربی برای یافتن تعداد صحیح امضاها استفاده می‌کند [۱۵].

---

[۱] biased

[۲] next-generation sequencers

[۳] isogenic

[۴] in vitro

[۵] Non-Negative Matrix Factorization

[۶] http://cancer.sanger.ac.uk/cosmic

از روش‌های دیگر، روشی براساس مدل تخصیص پنهان دیریکله[1] است [۱۶] که براساس مدل‌سازی موضوعی انجام می‌شود[۴]. در

مراجع

۱. Pfeifer GP, Denissenko MF, Olivier M, Tretyakova N, Hecht SS, Hainaut P. Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. Oncogene. 2002 Oct; 21 (48):7435–7451.

۲. Pfeifer GP, You YH, Besaratinia A. Mutations induced by ultraviolet light. Mutat Res. 2005 Apr; 571(1- 2):19–31.

۳. Stratton,M.R. et al. (2009) The cancer genome. Nature, 458, 719–724.

۴. Shiraishi,Y. et al. (2015) A simple model-based approach to inferring and visualizing cancer mutation signatures. PLoS Genet., 11, e1005657.

۵. Stratton,M.R. (2011) Exploring the genomes of cancer cells: progress and promise. Science, 331, 1553–1558.

۶. Harris,R.S. (2013) Cancer mutation signatures, dna damage mechanisms, and potential clinical implications. Genome Med., 5, 87.

۷. Temko,D. et al. (2018) The effects of mutational processes and selection on driver mutations across cancer types. Nat. Commun., 9, 1857.

۸. Wagener,R. et al. (2015) Analysis of mutational signatures in exomes from B-cell lymphoma cell lines suggest APOBEC3 family members to be involved in the pathogenesis of primary effusion lymphoma. Leukemia, 29, 1612–1615.

۹. Zou,X. et al. (2018) Validating the concept of mutational signatures with isogenic cell models. Nat. Commun., 9, 1744.

۱۰. Lee,D.D. and Seung,H.S. (2001) Algorithms for non-negative matrix factorization. In: Advances in Neural Information Processing Systems, pp. 556–562.

۱۱. Nik-Zainal,S. et al. (2012) Mutational processes molding the genomes of 21 breast cancers. Cell, 149, 979–993.

۱۲. Nik-Zainal,S. et al. (2016) Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature, 534, 47

۱۳. Alexandrov,L.B. et al. (2013a) Deciphering signatures of mutational processes operative in human cancer. Cell Rep., 3, 246–259.

۱۴. Alexandrov,L.B. et al. (2015) Clock-like mutational processes in human somatic cells. Nat. Genet., 47, 1402.

۱۵. Rosales,R.A. et al. (2017) Signer: an empirical bayesian approach to mutational signature discovery. Bioinformatics, 33, 8–16.

۱۶. Blei,D.M. et al. (2003) Latent dirichlet allocation. J. Mach. Learn. Res., 3, 993–1022.

۱۷.