

نامتعادلی لینکاژ (Linkage Disequilibrium)

گردآوری: رستم عبداللهی

نامتعادلی لینکاژ (LD) یکی از آن واژه‌هایی است که معنی خود را به خوبی آشکار نمی‌کند. همانطوری که هر فرد آشنایی با ژنتیک می‌داند، این واژه گنگ است و به فهم آن کمک نمی‌کند. LD به طور ساده جور شدن غیر تصادفی آلل‌ها در دو یا چند لوکوس می‌باشد و تعیین LD بود یا نبود قطعی لینکاژ را قطعی نمی‌کند. این واژه اولین بار در سال ۱۹۴۰ توسط Lowntin and kojima بکار رفت و استفاده از آن به همین منوال ادامه یافت. در ابتدا اطلاعات کمی برای مطالعه LD وجود داشت و اهمیت آن برای بیولوژیست‌های تکامل و ژنتیک دانان انسانی به استثناء ژنتیک جمعیت آشکار نبود. با این وجود، علاقه به LD به طور سریعی از سال ۱۹۸۰ به محض اطلاع از مفید بودن آن برای نقشه‌یابی ژن‌ها رشد پیدا کرد و تحقیقات زیادی برای شناسایی لوکوس‌های نزدیک به هم انجام گرفت. از آن پس این واژه آنقدر خوب در اذعان عموم پایه‌گذاری شد که دیگر نیازی به عوض کردن آن نبود.

LD هم در بیولوژی تکامل و هم در ژنتیک انسانی و غیر انسانی بسیار مهم است زیرا بر فاکتورهای تاثیر گذاشته و از فاکتورهای زیادی تاثیر می‌پذیرد. LD اطلاعاتی در ارتباط با رویدادهای گذشته فراهم می‌کند و پاسخ بالقوه به انتخاب طبیعی و مصنوعی را محدود می‌کند. LD سرتاسر ژنوم، تاریخ جمعیت را از نظر سیستم اصلاحی و الگوی تقسیم‌شدن جغرافیایی منعکس می‌کند. LD در هر منطقه ژنوم تاریخ انتخاب طبیعی، تغییر ژن، جهش، و سایر نیروهایی که باعث تکامل ژن- فراوانی می‌شوند را انعکاس می‌دهد. حال اینکه چگونه این فاکتورها بر LD بین جفت خاصی از لوکوس‌ها یا در منطقه ژنومی خاصی تاثیر می‌گذارند به نرخ نوترکیبی جایگاه بستگی دارد.

تعاریف

یک جفت لوکوس: LD بین آلل‌ها در دو لوکوس به طرق مختلفی تعریف شده است اما همه تعاریف به کمیت D بستگی دارند:

$$D_{AB} = P_{AB} - P_A P_B$$

که رابطه بالا تفاوت بین فراوانی گامت‌هایی که جفت آلل A و B را در دو لوکوس حمل می‌کنند (P_{AB}) و ضرب فراوانی آلل‌ها ($P_A P_B$) می‌باشد. کاربرد معمول این برای لوکوس‌هایی است که روی کروموزوم مشابهی هستند که در این مورد جفت آلل AB هاپلوتاوپ نامیده شده و P_{AB} فراوانی هاپلوتاوپ می‌باشد.

تبادل لینکاژ: اگر $D=0$ باشد تعادل لینکاژ برقرار است (LE). که به تعادل هاردی واینبرگ (HWE) شباهت دارد. زیرا بر این دلالت دارد که آلل‌ها در لوکوس‌های متفاوت به صورت تصادفی جور می‌شوند. فراوانی هاپلوتاوپ AB برابر با حاصلضرب $P_A P_B$ است. با این وجود LE از HWE متفاوت است. زیرا در یک نسل از آمیزش تصادفی برقرار نمی‌شود. در عوض، D بسته به فراوانی نوترکیبی (C) بین دو لوکوس بر اساس فرمول زیر کاهش می‌یابد.

$$D_{AB}(t + 1) = (1 - c)D_{AB}(t)$$

که در رابطه بالا t نشان دهنده نسل می‌باشد. حتی برای لوکوس‌های ناپیوسته $D (c=0/5)$ به اندازه $0/5$ در هر نسل کاهش می‌یابد. که این فرمول اولین بار توسط **Jenning** بدست آمد. اگرچه نهایتاً **LE** حاصل خواهد شد اما این اتفاق برای لوکوس‌های نزدیک به هم بسیار آهسته خواهد بود.

فاز هاپلوتایپ

D و آماره‌های مرتبط به سادگی فرض می‌کنند که افراد هاپلوئید و یا گامت‌ها می‌توانند شناسایی شوند. اما اغلب تنها ژنوتیپ‌های دیپلوئید می‌توانند تعیین شوند. اگر فردی دارای ژنوتیپ **AaBbCc** باشد این فرد می‌تواند هاپلوتایپ‌های **ABC/abc** و **ABC/abC** و **aBc/AbC** یا **aBC/Abc** را ایجاد کند. تعداد هاپلوتایپ ممکن به صورت نمایی با افزایش تعداد لوکوس هتروزیگوت افزایش می‌یابد. روش زیادی برای فائق شدن بر تشخیص هاپلوتایپ‌های یک فرد وجود دارد که می‌توان روش‌های بیوشیمیایی که به طور مجزایی هر کروموزوم را تکثیر می‌کنند اما این روش برای تحقیقات ژنومی در سطح وسیع گران و خیلی کم است. روش‌های دیگر شامل **Combinatorial theory**, **Parsimony**, **Metropolis algorithm based on ML** و توزیع پیشین بدست آمده از **Coalescent theory** می‌باشند.

LD در بیشتر از دو لوکوس

جعبه هاپلوتایپ: هنگامی که بیشتر از دو لوکوس با همدیگر در نظر گرفته می‌شوند یک عمل مهم اینست که از طریق گرافیکی آن جفت لوکوس‌هایی که سطوح بالایی از **LD** را دارند از آنهایی که مقدار کمتری **LD** دارند را تشخیص دهیم. که به این گراف **Haplotype Box** می‌گویند. جعبه هاپلوتایپ در انسان از نظر اندازه از فاصله یک کیلو بازی تا بیشتر از ۱۰۰ کیلو باز تغییر می‌کند. جعبه هاپلوتایپ کشف غافل گیر کننده‌ای بود که اهمیت عملی بسیار زیادی برای نقشه یابی بیماری‌ها وراثتی دارد.

واریانس در هتروزیگوتی: یک آماره ساده و اغلب مناسب برای بیان میزان کلی **LD** در یک منطقه ژنومی واریانس هتروزیگوسیته در میان لوکوس‌ها می‌باشد که به صورت یک تابع خطی از D^2 می‌باشد (Brown et al. 2005).

نامتعادلی رتبه بالاتر (High-order disequilibrium): هنگامی که بیشتر از دو لوکوس را در نظر می‌گیریم مقدار D می‌تواند به ضریب رتبه بالاتر **LD** تعمیم داده شود. برای آلل‌ها در سه لوکوس (**A**, **B** و **C**) ضریب رتبه ۳ به صورت زیر

$$D_{ABC} = P_{ABC} - P_A D_{BC} - P_B D_{AC} - P_C D_{AB} - P_A P_B P_C$$

که در رابطه بالا D_{BC} , D_{AC} و D_{AB} ضرایب نامتعادلی جفتی می‌باشند و D_{ABC} مشابه با اثر متقابل سه طرفه، در آنالیز واریانس می‌باشد و می‌تواند به عنوان نامستقلی در بین این لوکوس‌ها که به وسیله اثر متقابل دوتایی در نظر گرفته نمی‌شود تفسیر شود.

ژنتیک جمعیت LD

انتخاب طبیعی: علاقه اولیه به **LD** از سوالاتی در باره بکارگیری انتخاب طبیعی ایجاد می‌شود. اگر آلل‌های موجود در دو لوکوس در **LD** باشند و هر دو بر شایستگی تولید مثلی اثر کنند، پاسخ به انتخاب برای یک لوکوس ممکن است از طریق انتخاب برای لوکوس

دیگر کم یا زیاد شود. در برخی موارد انتخاب به تنهایی LD را افزایش می‌دهد. این مورد هنگامی است که فردی که هاپلوتایپ AB را حمل می‌کند میانگین شایستگی‌اش از افرادی که به تنهایی A یا B را حمل می‌کنند بیشتر باشد. اگر شایستگی نسبی (w) هاپلوتایپ‌های ab، Ab و AB به ترتیب ۱، w_{Ab} و w_{AB} باشد آنگاه اگر $w_{AB} > w_{Ab} * w_{Ab}$ باشد انتخاب LD را افزایش خواهد داد. پس همیشه اینطور نیست که انتخاب LD را افزایش دهد بلکه به شرایط بستگی دارد.

رانس ژنتیکی (Random Drift): رانس ژنتیکی تنها می‌تواند LD را بین لوکوس‌هایی که بسیار به هم نزدیک هستند ایجاد کند. اثر این شبیه به گرفتن یک نمونه کوچک از جمعیت بزرگ می‌باشد. رانس ژنتیکی همراه انتخاب به شیوه شگفت‌انگیزی برهمکنش دارد. انتخاب باعث می‌شود که لوکوس‌های به هم پیوسته خیلی نزدیک اندکی سست شوند زیرا رانس ژنتیکی باعث ایجاد مقدار کمی LD می‌شود که بر این اساس میانگین پاسخ به انتخاب کاهش می‌یابد. این اثر به اثر Hill-Robertson معروف است. هنگامی که تنها دو لوکوس در نظر گرفته می‌شوند ضعیف اما وقتی لوکوس‌های انتخاب شده زیادی به هم پیوسته و فاصله نزدیکی داشته باشند، بسیار قوی است.

تقسیمات جمعیتی و گلوگاه (bottleneck) جمعیت

انتخاب طبیعی تنها یک یا تعداد کمی از لوکوس‌ها را تحت تاثیر قرار می‌دهد. در مقابل تقسیمات جمعیتی، تغییر در اندازه جمعیت و مهاجرت افراد میان جمعیت‌ها LD را در سرتاسر ژنوم تحت تاثیر قرار می‌دهند. در نتیجه الگوی Genom-wide، نامتعادلی لینکاژ می‌تواند به ما در فهم تاریخ تغییر اندازه جمعیت و الگوی تبادل ژنی کمک کند.

مخلوط کردن عمدی یا غیر عمدی افراد از زیرجمعیت‌ها که فراوانی آلی متفاوت دارند باعث LD می‌شود. تغییر در اندازه جمعیت، به ویژه یک کاهش زیاد در اندازه (گلوگاه جمعیت) می‌تواند LD را افزایش دهد. بسیاری از مدل‌های تاریخ بشری فرض می‌کنند هنگامی که انسان‌های مدرن اولین بار آفریقا را ترک کردند یک گلوگاه رخ داده است. بعد از گلوگاه، برخی از هاپلوتایپ‌ها از بین خواهند رفت که عموماً منجر به افزایش LD خواهد شد. شناسایی سطوح بالای LD در سرتاسر ژنوم در یک جمعیت نسبت به جمعیت دیگر می‌تواند نشان از یک گلوگاه در گذشته باشد.

همخونی، وارونگی و تبدیل ژنومی

همخونی میزان LD را با چند برابر کردن کواریانس بین لوکوس‌ها در آلل‌های متفاوت افزایش می‌دهد. وارونگی ژنومی به مقدار زیادی نوترکیبی بین قطعات وارون شده و غیر وارون شده را کاهش می‌دهد زیرا که افراد نوترکیب آنوپلوئید هستند و قادر به تولید مثل نیستند.

کاربردهای LD

۱- جهش و نقشه‌یابی ژن‌ها: جهش یک نقش منحصر به فرد در ایجاد LD دارد. وقتی که در جایگاهی برای یک ژن دو آلل وجود داشته باشد و در جایگاه کناری آن فقط B وجود داشته باشد با جهش B به b، و افزایش فراوانی gamet ab و یا Ab

در اثر انتخاب باعث افزایش LD خواهد شد. تداوم LD قوی بین یک آلل جهش یافته و لوکوسی که بسیار به آن نزدیک است کاربردهای عملی بسیاری دارد.

۲- شناسایی انتخاب طبیعی: انتخاب مثبت قوی به طور سریعی فراوانی یک آلل مطلوب را افزایش می‌دهد و لوکوس‌های پیوسته به آن نیز به صورت غیر معمولی به صورت LD با آن باقی خواهند ماند. این ایده توسط میانارد اسمیت و های (۲۰۰۲) شکل گرفت ایشان این را Genetic hitch-hiking نامیدند. در مقاله ایشان تاکید شده که آلل مطلوب به سوی تثبیت شدن جلو می‌رود و باعث کاهش چشم‌گیر هتروزیگوسیتی در لوکوس‌های خنثی نزدیک به آن می‌شود. اخیراً، روش‌هایی برای شناسایی مناطقی از ژنوم که به صورت غیر عادی از میزان هتروزیگوسیتی آنها کاسته شده و نمادی از رویدادهای hitch-hiking می‌باشند، توسعه داده شده است.

۳- تخمین سن آلل: LD بسیار زیاد با یک آلل در یک منطقه نسبتاً وسیع نشان می‌دهد که زمان زیادی نمی‌گذرد که این الل به و سیله جهش ایجاد شده است. اگر آلل جهش یافته به یک فراوانی نسبتاً بالایی رسیده باشد نشان از این است که این آلل تحت تاثیر انتخاب مثبت قرار گرفته است.

کروم اسکن (CHROMSCAN): بسیاری از مطالعات Genomic-wide association mapping با استفاده از تابلوهای تراکم بالا (High-density panels) از SNP ها در حال حاضر قابل دسترس هستند. عموماً این‌ها شامل صدها تا هزاران نمونه DNA از حالت‌ها و مشاهده‌ها می‌باشد و برای چندین صدهزار SNP در میان ژنوم غربال می‌شوند. آنالیز موثر و تفسیر این مجموعه داده زیاد مشکلات بسیار زیادی دارد. که در این مورد برنامه CHROMSCAN می‌تواند بسیار مهم باشد. این برنامه از یک مدل درست‌نمایی ترکیبی (Composite-maximum likelihood model) برای آنالیز associton data فراهم می‌کند. مکان یابی ژن‌های بیماری بر اساس واحد نقشه نامتعادلی لینکاژ (LDU) مکان و خطای استاندارد برای چند شکلی‌های causative به وسیله این برنامه تعیین می‌شود.

TEAM: یک ابزاری برای ادغام بیان، لینکاژ و نقشه association می‌باشد.