

Trends and Future Prospects in Comparative Genomics as a Tool for Studying Evolution

Safari^{1*}, S. and Safari², Z.

Received Date: 13 September 2023

Accepted Date: 26 November 2023

(Review Article)

Abstract

Comparative genomics is a valuable tool for expanding molecular resources in crop plants, identifying key genes in their development, and studying the evolution of gene families. Tremendous advances in sequencing have opened up new avenues for comparative genomics, enabling researchers to perform large-scale genome-wide comparisons. Comparative genomics involves comparing any type of gene map at any level, from cytological to molecular. The main principle of comparative genomics is that shared features between two organisms should often be encoded in DNA that has been conserved between them during evolution. One of the major goals of this field is to identify the mechanisms of evolution of eukaryotic genomes. Identifying the locus of beneficial genes is an important step in breeding crops that are optimized for higher yield, quality, and disease resistance. This method is not only powerful but also rapid. As DNA sequencing technology has become more powerful and cheaper, comparative genomics has found wider practical applications in agriculture, biotechnology, and zoology as a tool for dissecting subtle differences between species. Such efforts are leading to new insights into some branches of the evolutionary tree, as well as improved health and new strategies for conserving rare and endangered species.

Keywords: Comparative genomics, Phylogenetics, Evolutionary tree

1. Researcher of Ilam Agricultural and Natural Research and Education Center, Ilam, Iran

2. Ph.D. Graduated, Department of Plant Breeding, University of Ilam, Ilam, Iran

*: Corresponding author Email: s_safari@areeo.ac.ir

روندها و چشم‌اندازهای آینده در ژنومیک مقایسه‌ای به‌عنوان ابزاری برای مطالعه تکامل

Trends and Future Prospects in Comparative Genomics as a Tool for Studying Evolution

سارا صفری^{۱*} و زینب صفری^۲

تاریخ پذیرش: ۱۴۰۲/۰۹/۰۵

تاریخ دریافت: ۱۴۰۲/۰۶/۲۲

(مقاله مروری)

چکیده

ژنومیک مقایسه‌ای ابزاری ارزشمند در گسترش منابع مولکولی برای گیاهان زراعی و شناسایی ژن‌های کلیدی در توسعه این گیاهان و مطالعه تکامل خانواده‌های ژنی است. پیشرفت‌های فوق‌العاده در توالی‌یابی، راهی جدید به ژنومیک مقایسه‌ای باز کرده و محققان را قادر به انجام مقایسات گسترده ژنومی ساخته است. ژنومیک مقایسه‌ای شامل مقایسه هر نوع نقشه ژنی در هر سطح از سیتولوژیکی تا مولکولی است. اصل عمده ژنومیک مقایسه‌ای این است که ویژگی‌های مشترک دو ارگانیسم اغلب باید در DNA یی که در طول تکامل بین آن‌ها حفظ شده، رمزگذاری شود. یکی از اهداف مهم این رشته، شناسایی مکانیسم تکامل ژنوم‌های یوکاریوتی است. شناسایی جایگاه ژن‌های سودمند گام مهمی در اصلاح محصولات زراعی است که برای عملکرد بیشتر، کیفیت، و مقاومت در برابر بیماری بهینه‌سازی می‌شوند. این روش نه تنها قدرتمند، بلکه سریع نیز هست. همان‌طور که تکنولوژی توالی‌یابی DNA قوی‌تر و ارزان‌تر شده است، ژنومیک مقایسه‌ای کاربردهای عملی گسترده‌تری را در بخش کشاورزی، بیوتکنولوژی و جانورشناسی به‌عنوان ابزاری برای جدا نمودن تفاوت‌های ظریف میان گونه‌ها فراهم کرده است. چنین تلاش‌هایی منجر به بینش‌های جدیدی در برخی از شاخه‌های درخت تکاملی و همچنین بهبود سلامت و با استراتژی‌های جدید برای حفظ گونه‌های نادر و در خطر انقراض است.

واژه‌های کلیدی: ژنومیک مقایسه‌ای، فیلوژنتیک، درخت تکاملی

۱. محقق، مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی استان ایلام، ایران

۲. دانش‌آموخته دکتری، گروه اصلاح نباتات، دانشگاه ایلام، ایلام، ایران

* نویسنده مسوول Email: s_safari@areeo.ac.ir

مبانی تاریخی و چارچوب مفهومی ژنومیک مقایسه‌ای

پیشرفت‌های انقلابی در فناوری توالی‌یابی که عمدتاً در پی پروژه ژنوم انسان شکل گرفت، تنها به تولید داده‌های ژنومی این پروژه محدود نشد. این دستاوردها با ایجاد تحول در روش‌های آزمایشگاهی، امکان تولید انبوه داده‌های توالی‌یابی را با هزینه‌ای به مراتب کمتر فراهم آورد. به لطف این تحولات فناورانه، توالی‌یابی ژنوم‌های متعدد مهره‌داران که برای درک و تفسیر ژنوم انسان ضروری هستند میسر شد. در نهایت، این توانمندی‌های نوین، بنیان مطالعات در مقیاس بزرگ را در حوزه‌هایی مانند تکامل ژنوم مهره‌داران، ژنومیک مقایسه‌ای و همچنین پزشکی انسانی تقویت کرد (Anderson et al., 2004). در گذر زمان، ژنومیک مقایسه‌ای نه تنها به عنوان ابزاری اساسی برای کشف مکانیسم‌های تکاملی، بلکه به عنوان عاملی کلیدی برای درک صحیح و تفسیر ژنوم انسان اثبات شده است. این حوزه سهم به سزایی در کشف مکانیسم‌های تکاملی گیاهان ایفا کرده است. با توجه به اندازه نسبتاً بزرگ و پیچیدگی ژنوم‌های گیاهی، انجام تحلیل‌های مقایسه‌ای و عملکردی اگرچه کاری دشوار، اما در عین حال بسیار ضروری و تعیین‌کننده به شمار می‌رود (Wang et al., 2023).

اصول بنیادین و اهمیت ژنومیک مقایسه‌ای در زیست‌شناسی تکاملی

ژنومیک مقایسه‌ای سهم به سزایی در روشن‌سازی بنیان ژنتیکی تفاوت‌های ظاهری میان گونه‌ها و حتی درون یک گونه دارد. با این حال، یک چالش مهم در مطالعات بین‌گونه‌ای، ناتوانی در تشخیص آن دسته از واریانت‌های آلی است که بر اثر جهش اضافه به وجود آمده و نقش تعیین‌کننده‌ای ایفا می‌کنند. شناسایی این واریانت‌ها عموماً تنها در حالتی امکان‌پذیر است که جهش در ناحیه‌ی کدکننده یک ژن رخ داده و عملکرد آن ژن نیز از پیش به خوبی شناخته شده باشد (Barcia-Cruz et al., 2024). مطالعه افرادی که فنوتیپ‌های افراطی در یک گونه دارند، می‌تواند به عنوان روشی مکمل برای مطالعات ژنومیک مقایسه‌ای بین‌گونه‌ای عمل کند. این نوع از ژنومیک مقایسه‌ای برای شناسایی توالی‌های غیرکدکننده‌ای که از نظر عملکردی حیاتی هستند، ابزاری ضروری است. برای نمونه، مقایسه ژنوم انسان و موش آشکار ساخته که حدود ۵ درصد از ژنوم پستانداران در طول تکامل حفاظت شده است. نکته جالب اینجاست که تنها ۱/۵ درصد از این ژنوم حفاظت‌شده، مسئول کد کردن پروتئین‌ها است و این موضوع نشان می‌دهد ۳/۵ درصد باقیمانده متعلق به مناطق غیرکدکننده‌ای است که اهمیت عملکردی دارند. در نتیجه، مقایسه ژنوم‌های گونه‌های

مختلف، احتمال شناسایی این توالی‌های حفاظت‌شده تکاملی را به‌طور قابل توجهی افزایش می‌دهد (Anderson et al., 2004). ژنومیک مقایسه‌ای ابزاری بنیادی برای بازسازی روابط تکاملی و ترسیم درخت فیلوژنتیک میان گونه‌ها و نژادهای مختلف است. این مقایسه‌ها می‌تواند در سطوح گوناگونی از نژادهای درون یک گونه تا گونه‌های دور از هم انجام پذیرد. معیارهای مقایسه ژنوم‌ها بسیار متنوع است و شامل مواردی همچون تعداد ژن‌ها، محتوای توالی، آرایش ژن‌ها روی کروموزوم، و نیز سازمان‌دهی خوشه‌ای ژن‌های مرتبط با یک مسیر متابولیکی می‌شود. یکی از کلیدی‌ترین معیارها، هم‌خطی است که به حفظ آرایش بلوک‌های ژنی در ژنوم‌های مختلف اشاره دارد. اهمیت آرایش ژن‌ها به حدی است که در گونه‌های بسیار نزدیک، این ترتیب به شدت حفاظت می‌شود، اما با افزایش فاصله تکاملی، بازآرایی‌های کروموزومی باعث از بین رفتن تدریجی هم‌راستایی ژن‌های ارتولوگ می‌شود. در نهایت، یک ژنوم کامل را می‌توان نقشه نهایی یک موجود دانست؛ نقشه‌ای که در آن تمامی صفات وراثتی در قالب توالی دقیق نوکلئوتیدهای هر کروموزوم رمزگذاری شده‌اند (Anderson et al., 2004).

مبنای اصلی ژنومیک مقایسه‌ای بر این اصل استوار است که اگر دو ارگانیسم یک ویژگی مشترک دارند، اطلاعات ژنتیکی مربوط به آن ویژگی می‌بایست در بخش‌هایی از DNA آن‌ها که در طول فرآیند تکامل حفظ شده‌اند، نهفته باشد (Hardison, 2003). بنابراین، روش‌های ژنومیک مقایسه‌ای با ساخت برخی از ترتیبات دلخواه توالی‌های ژنومی شروع و توالی‌های ارتولوگ (توالی‌هایی که یک جد مشترک را به اشتراک می‌گذارند) که در ژنوم‌های مرتب شده قرار دارد را جستجو می‌کند و چک می‌کند که تا چه حد این توالی حفاظت شده است. بر این اساس، ژنوم و تکامل مولکولی استنباط می‌شود و این ممکن است به نوبه خود به عنوان مثالی در محتوای تکامل فنوتیپی یا ژنتیک جمعیت قرار گیرد (Ellegren, 2008).

ژنومیک مقایسه‌ای با تحلیل هم‌زمان تشابهات و تمایزهای موجود در مولکول‌های پروتئینی، RNA و نواحی تنظیمی ژن در گونه‌های مختلف، به کشف نحوه انتخاب طبیعی اعمال شده بر این عناصر می‌پردازد. بر این اساس، عناصر ژنتیکی که مسئول ایجاد شباهت‌های بین‌گونه‌ای هستند، معمولاً تحت فشار انتخاب حفظ می‌شوند، درحالی‌که عناصری که باعث ایجاد تفاوت‌ها می‌شوند، طی فرآیند واگرایی تکاملی دست‌خوش تغییر می‌گردند. در نهایت، عناصری که نقش حیاتی در بقای موجود زنده ندارند، فاقد الگوی حفاظت تکاملی خواهند بود (Darling et al., 2008).

آنجا که الگوی سایت‌های برشی در کلون‌های هم‌پوشان به دقت مشخص می‌شود، می‌توان با کنار هم گذاشتن این اطلاعات، یک نقشه فیزیکی پیوسته (کانتیگ) از نواحی DNA ایجاد کرد (Zhou et al., 2024).

یکپارچه‌سازی و هم‌راستاسازی ژنوم‌ها

ارزش اصلی در قابلیت ارجاع متقابل بین نقشه‌های مختلف نهفته است، چرا که این ویژگی، گذار از نقشه ژنتیکی به نقشه فیزیکی و در نهایت به نقشه مولکولی را ممکن می‌سازد. این ارجاعات، بستری برای شناسایی و جداسازی فیزیکی ژن‌ها ایجاد می‌کنند. عامل محدودکننده در این فرآیند مقایسه‌ای، ماهیت نشانگرهای مورد استفاده در ترسیم هر نقشه است. به‌طوری‌که نقشه‌های ژنتیکی اساساً از نشانگرهای بسیار پلی‌مورف DNA تشکیل می‌شوند، اما نقشه‌های فیزیکی عمدتاً بر اساس ژن‌های کلون شده ساخته می‌شوند. بنابراین، یک استراتژی کاربردی، ادغام ژن‌های کدکننده در نقشه‌های ژنتیکی است تا امکان انتقال آن‌ها به نقشه‌های فیزیکی و انجام مقایسه‌های بین‌گونه‌ای فراهم آید (Flores et al., 2023).

سطوح مختلف وضوح در نقشه‌برداری ژنتیکی و فیزیکی به‌صورت سلسله‌مراتبی قابل تفکیک هستند. در یک سو، روش‌هایی با قدرت تفکیک پایین (حدود ۳۰ مگاباز) مانند هیبریداسیون در محل و ژنتیک سلول‌های سوماتیک قرار دارند و در سوی دیگر، نقشه‌برداری پیوندی با دقتی در حد سانتی‌مورگان (معادل ۱ تا ۲ مگاباز) واقع شده است. شکاف موجود بین این نقشه‌ها و سطح مولکولی (با حداکثر وضوح در حد توالی نوکلئوتیدی) از طریق کلون‌های درجی بزرگ (نظیر BACها و YACها) و ساخت کانتیگ‌های گسترده پر می‌شود. در این میان، هیبریدهای تابشی با فراهم آوردن وضوحی متوسط و امکان مرتب‌سازی قطعات، به‌عنوان چارچوبی برای اسمبل کانتیگ‌ها عمل می‌کنند. در سطح مولکولی، نقشه‌های برشی با نمایش آرایه‌ای خطی از سایت‌های برش آنزیم‌های محدودگر در فواصل چندکیلوبازی، دقتی بینابین را ارائه می‌دهند (Sengupta & Azad, 2023).

شکاف نقشه‌های مقایسه‌ای

نقشه‌های مقایسه‌ای موجود با شکاف‌های اطلاعاتی متعددی روبرو هستند و داده‌های مربوط به بسیاری از ژنوم‌های بالارزش، ناقص است. اگرچه در ابتدا توجهی به نقشه‌برداری ژنتیکی با سلول‌های سوماتیک وجود داشت، اما این روش تنها بر روی تعداد محدودی از گونه‌ها به‌طور کامل پیاده‌سازی شده است.

ژنومیک مقایسه‌ای در گذر زمان ثابت کرده است که نه تنها برای کشف مکانیسم‌ها و نیروهای تکاملی، بلکه برای درک بهتر و تفسیر خود ژنوم نیز ابزاری ارزشمند محسوب می‌شود. فرضیه بنیادین این حوزه بر این اصل استوار است که اگر توالی DNAیی در میان گونه‌های متعدد و دور از هم حفظ شده باشد، این حفظ‌شدگی به دلیل فشارهای انتخابی صورت گرفته و به‌طور ضمنی حاکی از یک عملکرد زیستی مهم است. با این حال، عکس این گزاره همیشه صادق نیست؛ یک توالی DNA می‌تواند عملکرد بیولوژیکی مشخصی داشته باشد، حتی اگر در ژنوم گونه‌های دیگر مشابهتی برای آن یافت نشود. این موضوع به‌ویژه در مورد تغییرات تکاملی جدید که منحصر به یک دودمان خاص هستند، مصداق دارد؛ چرا که زمان کافی برای ظهور نشانه‌های حفاظت در توالی‌های هم‌خط سپری نشده است. علاوه بر این، مفهوم "حفاظت" به معنای "همانی توالی" نیست، بلکه به معنای محدودیت در تغییر است؛ به‌طوری‌که یک توالی ممکن است در موقعیت‌های خاص خود، تنها تمایل به حفظ دو یا سه باز از چهار باز نوکلئوتیدی را داشته باشد (Alföldi et al., 2013).

از توالی‌یابی تا نقشه‌برداری: گردآوری داده‌های ژنومی

روش‌های نقشه‌برداری فیزیکی کلاسیک نظیر استفاده از تکنیک‌های ژنتیکی سلول‌های سوماتیک یا هیبریداسیون در محل، از محدودیت‌های چشمگیری در دقت رنج می‌برند. در تقابل با این روش‌ها، رویکردهای جایگزینی توسعه یافته‌اند که قادر به ارائه اطلاعات فیزیکی در سطح مولکولی با دقت بالا هستند. ازجمله این روش‌ها می‌توان به نقشه‌برداری برشی توالی‌های کلون شده، اسمبل کانتیگ و در نهایت تعیین توالی کامل ژنوم اشاره نمود. در روش نقشه‌برداری برشی، آنزیم‌های برشی (اندونوکلازهای محدودگر)، توالی‌های نوکلئوتیدی مشخص و معمولاً پالیندرومی (عموماً ۴ یا ۶ جفت بازی) را تشخیص داده و DNA را برش می‌دهند. قطعات حاصل که دارای ابعاد مولکولی مشخصی هستند، به وسیله تکنیک الکتروفورز تفکیک و قابل رؤیت می‌گردند. نهایتاً، توالی و آرایش فضایی سایت‌های برشی را می‌توان از طریق آنالیز تطبیقی الگوی باندهای حاصل از هضم آنزیمی به‌صورت تکی و ترکیبی استنباط نمود.

ایجاد نقشه‌های برشی با گستره وسیع‌تر نیز با به‌کارگیری آنزیم‌های برشی که جایگاه‌های شناسایی طولانی (۸ بازی، ۱۰ بازی و فراتر از آن) دارند، امکان‌پذیر است. در این حالت، می‌توان هر قطعه کلون‌شده DNA، از مقیاس کوچک تا قطعات مگابازی (مانند YACها)، را مورد تحلیل و مقایسه قرار داد. از

مشکل عمده، منحصربه‌فرد بودن مجموعه نشانگرها (مانند ریزماهواره‌ها یا پلی‌مورفیسم‌های تک‌نوکلئوتیدی) در هر مطالعه است. از آنجا که هر گروه تحقیقاتی از مجموعه نشانگرهای متفاوتی برای نقشه‌برداری استفاده می‌کند، تطبیق و مقایسه مستقیم نقشه‌های ژنتیکی بین گونه‌های مختلف غیرممکن یا بسیار دشوار می‌شود. این امر مانند مقایسه دو نقشه جغرافیایی است که یکی با واحد کیلومتر و دیگری با واحد مایل ترسیم شده باشد. در مطالعات پیوندی که برای ردیابی بیماری‌های ارثی در جمعیت‌ها حیاتی هستند، عدم وجود داده‌های شجره‌نامه‌ای دقیق و کامل یک مانع بزرگ است. بدون اطلاعات دقیق از روابط خویشاوندی بین افراد، ردیابی الگوی توارث آلل‌های بیماری‌زا با دقت کافی امکان‌پذیر نیست. با وجودی که کمبود داده برای گونه‌های مهم شگفت‌انگیز به نظر می‌رسد، این خلأ به وضوح نشان می‌دهد که جامعه علمی نیازمند یک راه‌حل استاندارد و یکپارچه است. توسعه و استفاده از مجموعه‌ای از نشانگرهای جهانی و پراکند، همراه با تسریع در توالی‌یابی ژنوم‌های مرجع برای گونه‌های مختلف، می‌تواند این شکافها را پر کرده و بنیان محکمی برای مقایسه‌های بین‌گونه‌ای فراهم کند (Goffeau et al., 1996).

اصول تکاملی

نظریه تکامل، شالوده نظری رشته ژنومیک مقایسه‌ای را تشکیل می‌دهد و در عین حال، خود این نظریه نیز به واسطه یافته‌های ژنومیک مقایسه‌ای، توسعه و غنای بیشتری یافته است. مقایسه توالی دو یا چند ژنوم، در اصل به‌منظور استنباط روابط تکاملی و ترسیم درخت فیلوژنتیکی انجام می‌گیرد. داده‌های گسترده ژنومی، امکان مطالعه عملکرد ژن‌ها را در مقیاس کل ژنوم و در چارچوب تکامل مولکولی فراهم ساخته‌اند. با استفاده از انواع اطلاعات حاصل از ژنوم‌های زیستی و نیز با بررسی روندهای تکاملی عمودی و افقی، می‌توان به درک بهتری از ساختار ژن‌ها و مکانیسم‌های تنظیمی آن‌ها دست یافت. با این حال، یک چالش مهم در این تحلیل‌ها وجود دارد: بین ۱/۵ تا ۱۴/۵ درصد از ژن‌ها در یک ژنوم، حاصل انتقال ژن جانبی (انتقال ژن بین گونه‌ها) هستند. در این موارد، تفاوت در توالی‌ها نه به دلیل تکامل عمودی، بلکه به دلیل این مهاجرت ژنی است. بنابراین، برای انجام یک تحلیل سیستماتیک قابل اعتماد، ضروری است که یک مدل تکاملی نسبتاً کامل طراحی گردد تا از تأثیر مخرب انتقال ژن جانبی و نیز فقدان داده‌های توالی از گونه‌های کلیدی که دارای توالی‌های حفاظت‌شده هستند، جلوگیری به‌عمل آید (Argos et al., 1984).

ژنومیک مقایسه‌ای یکی از قدرتمندترین ابزارهای زیست‌شناسی مدرن است که به ما امکان می‌دهد نه تنها تاریخچه تکاملی موجودات را بازسازی کنیم، بلکه عملکرد ژن‌های ناشناخته را نیز پیش‌بینی نماییم. نکته جالب این است که این رویکرد مبتنی بر یک منطق ساده اما عمیق است: تکامل گرچه تنوع بی‌پایانی ایجاد کرده، اما در عین حال آثار قابل ردیابی از خود به جا گذاشته است. با این حال، باید توجه داشت که هم‌خطی کامل بین گونه‌های حتی نزدیک نیز به ندرت دیده می‌شود. بازآرایی‌های ژنومی، حذف و اضافات، و رویدادهای تکاملی مستقل، همگی باعث می‌شوند که این شباهت‌ها پیچیده و گاهی ناقص باشند. بنابراین، درحالی‌که ژنومیک مقایسه‌ای ابزار ارزشمندی است، تفسیر نتایج آن نیازمند در نظر گرفتن این پیچیدگی‌هاست (Yang et al., 2019).

روندهای نوظهور در ژنومیک مقایسه‌ای

در سال‌های اخیر، حوزه ژنومیک مقایسه‌ای شاهد چند روند کلیدی بوده است که جهت‌گیری آینده این رشته را شکل می‌دهند.

گذار از ژنوم تکی به پان‌ژنوم (Pan-genomes)

تمرکز مطالعات از بررسی یک ژنوم مرجع به سمت تهیه و مقایسه پان‌ژنوم‌ها تغییر یافته است. پان‌ژنوم که شامل مجموعه کامل توالی‌های ژنی درون یک گونه می‌شود، تصویر بسیار کامل‌تری از تنوع ژنتیکی و مسیرهای تکاملی ارائه می‌دهد (Golics et al., 2016).

ادغام داده‌های چند-آمیکی (Multi-omics)

روند غالب دیگر، تلفیق داده‌های حاصل از حوزه‌های مختلف زیست‌شناسی مولکولی ژنومیک مانند ژنومیک، ترانسکریپتومیکس، پروتئومیکس و اپی‌ژنومیکس است. این رویکرد یکپارچه، امکان درک مکانیسم‌های عملکردی پشت تغییرات توالی را فراهم می‌سازد (Yugi et al., 2016).

تسخیر هوش مصنوعی و یادگیری ماشین

با حجم شدن داده‌های ژنومی، استفاده از الگوریتم‌های هوش مصنوعی برای شناسایی الگوهای پیچیده تکاملی، پیش‌بینی عملکرد ژن‌ها و آنالیز هم‌راستایی‌ها در مقیاس بزرگ، به یک روند اجتناب‌ناپذیر تبدیل شده است (Avsec et al., 2024).

کاربردهای مستقیم در انتخاب مصنوعی و پزشکی

رشد فزاینده‌ای به سمت استفاده عملی از ژنومیک مقایسه‌ای برای شتاب بخشیدن به برنامه‌های به‌نژادی محصولات زراعی و دامی و نیز شناسایی هدف‌های دارویی جدید در پزشکی در حال وقوع است (Varshney *et al.*, 2020).

الگوریتم‌های هم‌راستایی و آنالیز فیلوژنتیک

رشد انفجاری داده‌های ژنومی در دهه‌های اخیر، محرک اصلی توسعه شتابان ابزارهای محاسباتی برای تحلیل توالی‌ها و ژنوم‌های کامل بوده است. این پیشرفت‌ها به موازات تکامل قابل توجهی در روش‌های تحلیل مقایسه‌ای صورت پذیرفته است. در این میان، یکی از چالش‌های محوری، بازنمایی مؤثر و تصویرسازی نتایج حاصل از مقایسه‌های ژنومی است.

جدول ۱. خلاصه‌ای از ابزارهای نرم‌افزاری پرکاربرد در ژنومیک مقایسه‌ای

Table 1. Summary of commonly used software tools in comparative genomics

نام ابزار Tool name	کاربرد اصلی Main usage	دسترسی Accessibility
BLAST	یافتن توالی‌های مشابه در پایگاه‌های داده به‌عنوان سنگ بنای بسیاری از تحلیل‌های مقایسه‌ای	https://blast.ncbi.nlm.nih.gov/
UCSC Genome Browser	مرور بصری و مقایسه ژنوم‌ها در کنار انواع آنوتاسیون‌ها (شناسایی توالی‌های تنظیمی)	https://genome.ucsc.edu/
Ensembl	مرور، آنالیز و مقایسه ژنوم‌های به ویژه مهره‌داران و ارائه ابزارهای قدرتمند برای مقایسه هم‌ترازی	https://www.ensembl.org/
MUMmer	هم‌راستایی سریع و دقیق ژنوم‌های کامل بزرگ و ایده‌آل برای مقایسه ژنوم‌های باکتریایی یا یافتن تغییرات ساختاری	https://mummer4.github.io/
OrthoFinder	استنتاج ژن‌های ارتولوگ (هم‌تکامل) و پارالوگ در مقیاس ژنوم و اساس بسیاری از مطالعات فیلوژنومی	https://github.com/davidemms/OrthoFinder
MAUVE	هم‌راستایی ژنوم‌های چندگانه با در نظر گرفتن بازآرایی‌های ژنومی، عالی برای مقایسه گونه‌های نزدیک	http://darlinglab.org/mauve
VISTA	یک پلتفرم برای آنالیز حفاظت توالی‌های ژنومی و شناسایی عناصر تنظیمی حفاظت‌شده	https://genome.lbl.gov/vista
Circos	تولید نمودارهای حلقوی (Circular Plots) برای نمایش بصری مقایسه‌های ژنومی (مانند سینتینی)	http://circos.ca/

مزیت استفاده از ابزارهای آنلاین این است که این وب‌سایت‌ها در حال توسعه هستند و به‌طور مداوم به‌روز می‌شوند. بسیاری از تنظیمات جدید و مطالب وجود دارد که می‌توان آن‌ها را برای بهبود بهره‌وری به‌طور آنلاین مورد استفاده قرار داد (Haft, 2015).

کاربردهای ژنومیک مقایسه‌ای

کشاورزی

کشاورزی مزرعه‌ای است که مزایای ژنومیک مقایسه‌ای را درو می‌کند. کشاورزی مدرن به‌طور فزاینده‌ای در حال بهره‌برداری

از مزایای ژنومیک مقایسه‌ای است تا چالش‌های جهانی امنیت غذایی را پاسخ دهد. هدف غایی این تلاش‌ها، شناسایی ژن‌های کلیدی و واریانت‌های آللی است که صفات اقتصادی مهم (QTLs) از قبیل عملکرد، بازدهی، کیفیت محصول، مقاومت به بیماری‌ها و تحمل به تنش‌های محیطی مانند خشکی و شوری را کنترل می‌کنند (Safari & Mehrabi, 2020; Tahmasebi *et al.*, 2022). روش‌های سنتی نقشه‌یابی ژنتیکی، مانند استفاده از نشانگرهای مولکولی اولیه، مستلزم ایجاد جمعیت‌های اصلاحی بزرگ و پیگیری توارث صفات در چندین نسل بود که فرآیندی زمان‌بر و پرهزینه محسوب می‌شد. در مقابل، ژنومیک

تکامل ژنوم غلات که قبلاً با استفاده از ابزارهای متعارف تحقیقات سیتوژنتیکی حل نشده بود کمک کرد. بنابراین مطالعات دقیق در زمینه ژنومیک مقایسه‌ای محصولات گیاهی باعث تسهیل استفاده از اطلاعات تولید شده از ژنوم یکی از غلات در بقیه شده است. این جنبه‌های مرتبط با نقشه‌برداری مقایسه‌ای و ژنتیکی را در بر می‌گیرد (Ellegren, 2008). در نتیجه، ژنومیک مقایسه‌ای با ارائه یک روش قدرتمند، سریع و دقیق، انقلابی در به‌نژادی محصولات ایجاد کرده و مسیر را برای توسعه نسل بعدی ارقام پرمحصول، مقاوم و پایدار هموار ساخته است.

پزشکی

در زمینه پزشکی نیز از مطالعات ژنومیک مقایسه‌ای سود می‌برند. واکسن‌شناسی به طور خاص پیشرفت‌های مفیدی را با توجه به مشکلات روش‌های ژنومی تجربه کرده است. در یک رویکرد شناخته شده به‌عنوان واکسن‌شناسی معکوس، محققان می‌توانند آنتی‌ژن کاندید برای توسعه واکسن را با تجزیه و تحلیل ژنوم یک عامل بیماری‌زا یا یک خانواده از عوامل بیماری‌زا کشف نمایند. با استفاده از یک رویکرد ژنومیک مقایسه‌ای با تجزیه و تحلیل ژنوم چند پاتوژن مرتبط می‌توان منجر به توسعه واکسن‌هایی شد که حافظت چندگانه دارند (Wienberg & Stanyon, 1998).

تحقیقات

ژنومیک مقایسه‌ای راه‌های جدیدی را در سایر بخش‌های پژوهش نیز بازی می‌کند. همان‌طور که تکنولوژی توالی‌یابی DNA در دسترس‌تر شده است، تعداد توالی‌های ژنوم رشد کرده است. با افزایش مخزن داده‌های ژنومی موجود، قدرت استنتاج ژنومی مقایسه‌ای نیز بیشتر شده و رشد کرده است. این روش به محققان اجازه داده تا اطلاعات تنوع ژنتیکی، بیان ژن و دینامیک تکاملی در پستانداران را جمع‌آوری نموده که با استفاده از داده‌های روش‌های قبلی غیرقابل تشخیص بودند (Wienberg & Stanyon, 1998).

تحلیل و مقایسه انتقادی یافته‌های اخیر در ژنومیک مقایسه‌ای
در سال‌های اخیر، مطالعات ژنومیک مقایسه‌ای با تمرکز بر پرسش‌های تکاملی عمیق‌تر و با بهره‌گیری از فناوری‌های جدید، شتاب بی‌سابقه‌ای داشته است. درک نقاط قوت، تضادها و محدودیت‌های این یافته‌ها برای ترسیم نقشه راه آینده این حوزه ضروری است. در ادامه، مهم‌ترین دستاوردهای پژوهشی اخیر در چند محور کلیدی به‌صورت انتقادی مقایسه شده‌اند.

مقایسه‌ای با فراهم آوردن چارچوبی برای مقایسه مستقیم ژنوم‌های گونه‌های مختلف و واریته‌های یک گونه، این نیاز را مرتفع ساخته است. این رویکرد، محققان را قادر می‌سازد تا به سرعت نواحی ژنومی حفاظت‌شده‌ای را که با صفات مطلوب مرتبط هستند، شناسایی کنند (Flores *et al.*, 2023).

در دسترس بودن نقشه مولکولی برای ژنوم همه غلات مهم، مطالعات انجام شده در نقشه‌برداری مقایسه‌ای را تسهیل می‌کند. مطالعات مقایسه‌ای حفاظت خوبی از نشانگرها در قطعات بزرگ کروموزومی ژنوم غلات را نشان داده است. با این حال، نقشه‌برداری از ژن‌های مقاومت دارویی (RGAها) جدا شده از برنج، جو و ارزن دمروباهی اورتولوژی محدودی نشان داد. در سال‌های اخیر، جداسازی و توالی‌یابی قطعات بزرگ DNA ژنومی (۵۰۰-۱۰۰۰ KB) از بسیاری از غلات و حبوبات، اطلاعات بیشتری در مورد حفاظت ترتیب ژنی بین ژنوم‌های مختلف غلات در سطح زیر مگاباز ارائه داده است (Chain *et al.*, 2003; Touchman, 2010). مطالعات ژنومیک مقایسه‌ای تعداد زیادی از دو برابر شدگی‌ها، جابجایی‌ها و وارونگی‌ها را نیز حل و فصل نمود و به تکامل ژنوم غلات که قبلاً با استفاده از ابزارهای متعارف تحقیقات سیتوژنتیکی حل نشده بود کمک کرد. بنابراین مطالعات دقیق در زمینه ژنومیک مقایسه‌ای محصولات گیاهی باعث تسهیل استفاده از اطلاعات تولید شده از ژنوم یکی از غلات در بقیه شده است. این جنبه‌های مرتبط با نقشه‌برداری مقایسه‌ای و ژنتیکی را در بر می‌گیرد (Ellegren, 2008). به‌طور خلاصه تأثیر کاربردهای ژنومیک مقایسه‌ای در کشاورزی را می‌توان موارد زیر دانست.

۱. **کشف ژن‌های کاندیدای مسئول صفات:** با مقایسه ژنوم گیاهان مقاوم و حساس به یک بیماری، می‌توان ژن‌های منحصر به فرد یا واریانت‌های خاصی را که پایه مقاومت هستند، شناسایی نمود.

۲. **بهینه‌سازی برنامه‌های به‌نژادی:** از طریق ردیابی واریانت‌های آللی مفید در جمعیت‌های وسیع با استفاده از نشانگرهای مبتنی بر توالی (GBS: Genotyping-By-Sequencing)، روند انتخاب ژنومی (Genomic Selection) شتاب چشمگیری یافته است.

۳. **بازگردانی صفات از گونه‌های وحشی:** بسیاری از ژن‌های مفید برای تحمل به تنش، در خویشاوندان وحشی محصولات کشاورزی حفظ شده‌اند. ژنومیک مقایسه‌ای مسیر دقیقی برای شناسایی و انتقال هدفمند این ژن‌ها به ارقام رایج فراهم می‌کند.

مطالعات ژنومیک مقایسه‌ای تعداد زیادی از دو برابر شدگی‌ها، جابجایی‌ها و وارونگی‌ها را نیز حل و فصل نمود و به

مقابل، تحلیل‌های صرفاً ژنومیکی، اگرچه می‌توانند نواحی حفاظت‌شده یا تحت انتخاب را با دقت بالا شناسایی کنند، اما در ارائه توضیح چرایی و چگونگی عملکرد این نواحی غالباً ناتوان هستند.

تحلیل‌های مبتنی بر هوش مصنوعی در مقایسه با روش‌های سنتی تر بیوانفورماتیکی

پژوهش‌های پیشرفته، مانند کار آوسک (Avsec) و همکاران (۲۰۲۱) نشان داده‌اند که الگوریتم‌های یادگیری عمیق می‌توانند از روی توالی DNA به تنهایی، الگوهای تنظیم ژن از فواصل بسیار طولانی را پیش‌بینی کنند. این توانایی، درک ما را از پیچیدگی تنظیم ژن متحول ساخته است. در عین حال ابزارهای کلاسیک مانند BLAST و OrthoFinder در شناسایی هم‌توالی‌ها و ژن‌های ارتولوگ بین گونه‌ها بسیار قدرتمند هستند، اما معمولاً فاقد توانایی پیش‌بینی پیامدهای عملکردی این هم‌توالی‌ها می‌باشند. این تفاوت روش‌شناختی منجر به یک شکاف فزاینده بین سرعت تولید داده و سرعت تفسیر بیولوژیکی آن شده است. در حالی که یک ابزار سنتی ممکن است یک نوکلئوتید حفاظت‌شده را شناسایی کند، یک مدل هوش مصنوعی می‌تواند پیش‌بینی کند که جهش در آن نوکلئوتید خاص چگونه بر اتصال فاکتور رونویسی و در نهایت بیان یک ژن هدف در بافت خاصی تأثیر می‌گذارد. این امر، مطالعات ژنومیک مقایسه‌ای را از مرحله توصیف هم‌توالی‌ها به سمت پیش‌بینی پیامدهای عملکردی سوق داده است.

توانایی‌ها، مزیت‌ها و چالش‌های ژنومیک مقایسه‌ای

از منظر روش‌شناختی، می‌توان تحول در این حوزه را در محورهای زیر ترسیم نمود:

توسعه الگوریتم‌های محاسباتی

افزایش کمی داده‌های ژنومی مستلزم ایجاد الگوریتم‌های کارآمد برای هم‌ترازی توالی‌ها، شناسایی نواحی حفاظت‌شده، و پیش‌بینی عناصر عملکردی است. تسریع فرآیند تشخیص مبتنی بر DNA مستلزم آگاهی جامع از قابلیت‌ها و محدودیت‌های ابزارهای مختلف ژنومیک مقایسه‌ای است. در این زمینه، تعیین توالی دقیق ژن‌ها از اهمیت بسزایی برخوردار است، به‌ویژه در حوزه‌هایی نظیر شناسایی پاتوژن‌ها. باید در نظر داشت که تحلیل‌های عملکرد ژن، مکانیسم‌های تنظیم ژن، برهم‌کنش‌های ژنی و مطالعات فیلوژنتیک، همگی وابسته به دستیابی به داده‌های صحیح از توالی نوکلئوتیدی و اسیدهای آمینه هستند. در نهایت، به‌کارگیری ابزارهای تحلیلی مناسب

مطالعات مبتنی بر پان‌ژنوم در مقایسه با مطالعات مبتنی بر ژنوم مرجع

پژوهش‌های پیشگامانه‌ای مانند مطالعه گولیکز (Golick) و همکاران (۲۰۲۰) بر روی گونه‌های گیاهی *Brassica oleracea* به وضوح نشان داده‌اند که ژنوم مرجع (Reference Genome) واحد که یک توالی ژنومی دیجیتالی است که به عنوان یک نماینده استاندارد برای مجموعه ژنی یک گونه خاص در نظر گرفته می‌شود. این ژنوم معمولاً از توالی‌یابی و مونتاژ ژنوم چندین فرد به‌دست می‌آید و به‌عنوان یک داربست یا نقشه پایه برای مقایسه توالی‌های جدید دیگر افراد همان گونه استفاده می‌شود، قادر به بازنمایی تمام تنوع ژنتیکی درون یک گونه نیست. این مطالعات تأکید دارند که بخش قابل‌توجهی از ژنوم اختیاری (Dispensable)، حاوی ژن‌های کلیدی مرتبط با سازگاری به عوامل محیطی و بیماری‌ها است. در مقابل، بسیاری از تحلیل‌های ژنومیک مقایسه‌ای کلاسیک‌تر، هنوز بر هم‌راستایی با یک ژنوم مرجع واحد متکی هستند. اگرچه این رویکرد برای مقایسه‌های اولیه بین گونه‌های دور از هم کارآمد است، اما اغلب قادر به شناسایی واریانت‌های ساختاری مهم و حضور و غیاب ژن‌ها (PAV: Presence-Absence Variation) نیست.

این تفاوت در مقیاس مطالعه، منبع مستقیمی برای تفاوت در نتایج است. برای مثال، یک ژن که در مطالعه‌ای مبتنی بر پان-ژنوم به‌عنوان یک ژن حفظ شده (Conserved Gene) و حیاتی شناسایی می‌شود، ممکن است در مطالعه‌ای دیگر که بر ژنوم مرجع واحد متمرکز است، کاملاً نادیده گرفته شود، زیرا این ژن در آن ژنوم مرجع خاص وجود ندارد. بنابراین، تعمیم‌دهی نتایج حاصل از مطالعات پان-ژنوم به کل یک گونه معتبرتر است، در حالی که یافته‌های مطالعات مبتنی بر ژنوم تکی می‌تواند تا حدی گمراه‌کننده باشد. این نشان می‌دهد انتخاب مقیاس مطالعه مستقیماً بر درک ما از اندازه واقعی تنوع ژنتیکی و مکانیسم‌های سازش تأثیر می‌گذارد.

ادغام داده‌های چند آمیکی در مقابل ژنومیک تکی: همانطور که یوگی (Yugi) و همکاران (۲۰۱۶) در مرور خود بر ترنس-أمیکس توضیح داده‌اند، ادغام داده‌های ژنومیک با لایه‌های اطلاعاتی دیگر مانند ترانسکریپتومیکس و پروتئومیکس، این توانایی را ایجاد می‌کند که نه تنها وجود یک واریانت ژنتیکی، بلکه تأثیر عملکردی آن بر بیان ژن، تولید پروتئین و نهایتاً فنوتیپ نیز بررسی شود. برای نمونه، این رویکرد می‌تواند نشان دهد که یک توالی تنظیمی حفاظت‌شده چگونه در پاسخ به یک محرک محیطی، بیان چندین ژن را هم‌زمان کنترل می‌کند. در

تأثیر مستقیمی بر دقت و کیفیت یافته‌های پژوهشی دارد (Flores *et al.*, 2023).

ظرفیت تولید داده‌های توالی‌یابی در مراکز جهانی طی دهه‌های اخیر از رشدی نمایی برخوردار بوده است. در هسته تمامی این فرآیندها، مفهوم هم‌خطی قرار دارد که اساس تمامی روش‌های هم‌ترازی کنونی را تشکیل می‌دهد. در حوزه ژنومیک مقایسه‌ای که هم‌زمان هم‌ترازی توالی‌ها و مقایسه ساختار ژنوم‌ها را در بر می‌گیرد، از روش‌های تخصصی برای شناسایی نواحی هم‌خط استفاده می‌شود. این نواحی شناسایی شده سپس به‌صورت مجزا مورد هم‌ترازی قرار می‌گیرند (Hardison *et al.*, 1997؛ Lund *et al.*, 2000). البته کاربرد ژنومیک مقایسه‌ای بین گونه‌ای در شناسایی چنین نواحی به مقدار زیادی تحت تأثیر فاصله تکامل بین گونه‌های مورد سوال می‌باشد. استفاده از مقایسات مبتنی بر ترازبندی برای کشف عناصر کاربردی حفاظت شده اصطلاحاً انگشت‌نگاری فیلوژنتیک گفته می‌شود (Gottgens *et al.*, 2000). این روش نیاز به یکسری اطلاعات قبلی از یک موتیف توالی پروتیینی را مرتفع نموده و مکملی برای آنالیز الگوریتمی فراهم می‌کند. مقایسات ژنومی در گونه‌های مرتبط نیز می‌تواند به تعیین اساس ژنتیکی تنوع فنوتیپی کمک کند و نواحی مخصوص گونه‌ای را که می‌تواند برای شناسایی هدف قرار گیرد آشکار سازد. این اثرات DNA ممکن است چند شکلی‌های تک نوکلئوتیدی (SNPها)، تعداد متفاوت تکرارهای پشت سرهم چندمکانی (MLVA: Multi-locus Variable-number Tandem Repeat Analysis) و تنوعات مشابه را به‌کار گیرد آرایه مربوط به هدف تحقیقی متفاوت، نیاز به دسترسی به یک ابزار ژنومیک مقایسه‌ای تطبیق‌پذیر را برجسته می‌کند (Read *et al.*, 2000).

چالش‌های تصویرسازی داده‌ها

ارائه گرافیکی روابط تکاملی، نواحی هم‌خطی، و الگوهای حفاظت توالی به‌عنوان یک ضرورت درآمده است، چرا که درک روابط پیچیده ژنومی بدون بازنمایی بصری به‌طور قابل‌توجهی با مشکل مواجه می‌شود. فرآیند تصویرسازی داده‌های حاصل از هم‌ترازی ژنومی، چالشی پیچیده‌تر از آنچه به نظر می‌رسد را نمایان می‌سازد. خروجی استاندارد ابزارهای هم‌ترازی عمدتاً در قالب فایل‌های متنی ارائه می‌شود که بازه‌های باقیمانده هم‌تراز شده را ثبت می‌نمایند. هرچند برخی نرم‌افزارها امکان خروجی با فرمت XML را نیز فراهم می‌کنند که تبادل داده با سایر ابزارها را تسهیل می‌بخشد. با این وجود، صرف‌نظر از قالب خروجی، حجم انبوه اطلاعات استخراج شده از ژنوم‌های مورد مقایسه، تفسیر مستقیم نتایج را با دشواری مواجه می‌سازد. از

این‌رو توسعه ابزارهای تصویرسازی کارآمد که قادر به مدیریت پیچیدگی این داده‌ها باشند، برای انجام مقایسه‌های جامع و قابل درک، امری ضروری است. این ابزارها باید امکان یکپارچه‌سازی داده‌های متنوع از جمله همولوژی‌های پایگاه‌داده، پیش‌بینی‌های ژنی و همچنین شناسایی عناصر تکراری، نقاط شروع و اتصال، و محل اتصال پروتئین‌ها را فراهم آورده و در نهایت به زیست‌شناسان در تحلیل‌های خود کمک نمایند. در این میان، چالش دیگر مربوط به محدودیت‌های ورودی این نرم‌افزارهاست که عموماً تنها از فرمت‌های خاصی پشتیبانی می‌کنند. بسیاری از این موضوعات به‌وسیله تعدادی از پیشرفت‌های اخیر در برنامه‌های تصویرسازی ترازبندی مقایسه‌ای ژنوم مثل PipMaker، Alfresco، Intronerator، VISTA، SynPlot و ACT مطرح می‌شوند (Gottgens *et al.*, 2001).

یکپارچه‌سازی داده‌ها از طریق مرورگرهای ژنومی

ابزارهایی مانند UCSC Genome Browser و Ensembl با فراهم آوردن لایه‌های متعدد اطلاعاتی از نشانگرهای ژنتیکی تا عناصر تنظیمی، امکان درک همه‌جانبه از معماری ژنوم را فراهم ساخته‌اند. روش‌های فیلوژنتیک کلاسیک نمی‌توانند به‌صورت قابل اطمینانی قابلیت طراحی قطعات تشخیصی DNA را که قادر به شناسایی تمام اعضای یک خانواده ژنی در یک گونه یا تمام سویه‌های یک گونه باشند، پیش‌بینی نمایند. بر مبنای چنین مقایسه‌های ژنومی، ابزارهای متعددی برای هم‌ترازی توالی و اتصال زیررشته‌ها توسعه یافته‌اند که برخی از آن‌ها به‌طور ویژه برای غلبه بر چالش‌های ناشی از پردازش توالی‌های طولانی نظیر ژنوم‌ها یا کروموزوم‌های کامل طراحی شده‌اند. با این حال، مسئله پیچیده‌تر، مدیریت و نمایش این حجم از داده‌هاست، چرا که هم‌ترازی‌های تولید شده و همچنین تفسیر و بصری‌سازی آن‌ها غالباً با دشواری‌های عملیاتی مواجه هستند (Alm *et al.*, 1999؛ Paulsen *et al.*, 2000).

ابزارهای ترازبندی ژنوم معمولاً با یک هدف مشخص طراحی شده‌اند: بعضی به راحتی به پیدا کردن هر شباهتی بین دو ژنوم دست می‌یابند و بقیه مخصوصاً توالی‌های کد کننده (مثل اگزون‌ها) را هدف قرار می‌دهند و ترتیب اگزون‌ها بین دو گونه متفاوت حفاظت شده است، بقیه روی نواحی بین ژنی و اینترونی تمرکز دارند تا سیگنال‌های تنظیمی حفاظت شده را تشخیص دهند. برخی مشکلات اصلی مرتبط با این اهداف در اندازه‌گیری بازآرایی‌ها (مثل سر خوردن اگزون‌ها یا سایر نواحی غیرهم‌خط بدست آمده از نو ترکیبی‌های درون مولکولی)، اضافه و حذف‌شدگی‌های بلند (توالی‌هایی که چندین ناحیه شباهت موضعی جدا شده به‌وسیله نواحی غیرمرتبط)، عناصر تکرار شده

تکاملی، ساختار ژنوم و پایه‌های مولکولی بیماری‌ها را متحول سازد.

ژنومیک مقایسه‌ای یک میدان هیجان‌انگیز از تحقیقات بیولوژیکی است که در آن محقق با استفاده از انواع ابزارها، از جمله آنالیز مبتنی بر کامپیوتر، به مقایسه توالی ژنوم کامل از گونه‌های مختلف می‌پردازد. با ویژگی‌های مقایسه‌ای دقیق که موجودات مختلف از جمله ژنوم موجودات از انسان به شامپانزه و به مخمر را تعریف می‌کنند، محققان می‌توانند مناطق تشابه و تفاوت را با دقت پیدا کنند. این اطلاعات می‌تواند به دانشمندان کمک کند تا درک بهتری از ساختار و عملکرد ژن‌های انسان و توسعه استراتژی‌های جدید برای مبارزه با بیماری‌های انسانی داشته باشند. شناسایی توالی DNA حفاظت شده، که در بسیاری از موجودات مختلف در طول میلیون‌ها سال حفظ شده، گام مهمی به سوی درک خود ژنوم است. این ژن‌های ضروری برای حیات و سیگنال‌های ژنومی که عملکرد ژن در بسیاری از گونه‌های برجسته کنترل می‌کنند را مشخص می‌کند. این به ما در درک بیشتر ژن‌های مرتبط با سیستم‌های مختلف بیولوژیکی کمک می‌کند.

ژنومیک مقایسه‌ای ابزار قدرتمندی برای مطالعه تکامل نیز فراهم می‌کند. با بهره‌گیری از آنالیز روابط تکاملی بین گونه‌ها و تفاوت‌های مربوطه در DNA آن‌ها، دانشمندان بهتر می‌توانند تغییرات ظاهر، رفتار و زیست‌شناسی موجودات زنده را در طول زمان شناسایی کنند. همان‌طور که تکنولوژی توالی‌یابی DNA قوی‌تر و ارزان‌تر شده است، ژنومیک مقایسه‌ای نیز کاربردهای عملی گسترده‌تری را در بخش کشاورزی، بیوتکنولوژی و جانورشناسی به‌عنوان ابزاری برای جدا نمودن تفاوت‌های اغلب ظریف میان گونه‌ها فراهم کرده است. چنین تلاش‌هایی منجر به بینش‌های جدیدی در برخی از شاخه‌های درخت تکاملی و همچنین بهبود سلامت حیوانات اهلی و با اشاره به استراتژی‌های جدید برای حفظ گونه‌های نادر و در خطر انقراض است.

ژنومیک مقایسه‌ای نتایج چشمگیری داشته است. محققان به‌طور فزاینده‌ای با استفاده از ژنومیک مقایسه‌ای به کشف مناطق مسئول توسعه رفتار انسان و سوخت و ساز بدن و استعداد ابتلا به بیماری دست یافته‌اند. این مطالعات در حال کشف مسیر ژن‌های مرتبط با بخش عصبی و رشد و نمو است. برخی از محققان با استفاده از ژنومیک مقایسه‌ای به فاش کردن پایه‌های ژنومی بسیاری از بیماری‌ها در حیوانات، با امید به‌دست آوردن بینش جدید در توسعه بیماری در انسان پرداخته‌اند. به‌طور کلی و بر اساس موارد مطرح شده می‌توان توانایی‌ها و قابلیت‌های این تکنیک را شامل موارد ایجاد

(مثل ژن‌های دو برابر شده، ترنسپوزون‌ها، SINE، LINE،ها و ...)، تکرارهای پشت سر هم و مشکلات ذاتی عناصر تنظیم ژن، شامل اندازه کوچک آن‌ها و مقاومت مرتبط با حذف و اضافه‌های کوچک یا جایگزینی‌ها را به اشتراک می‌گذارند. انواع دیگر برنامه‌های مورد استفاده شامل DIALIGN، ASSIRC، LSH-، WABA، GLASS، PipMaker/Blastz، MUMmer، ALL-PAIRS، Vmatch و MGA می‌باشد (Chain et al., 2003). تعداد زیادی از این برنامه‌ها مبتنی بر یک الگوریتم مؤثر مورد استفاده توسط Dumas و Nino در ۱۹۸۲ هستند که اتصالات دقیق به طول K به نام K-mer را پیدا می‌کند. هر K-mer، برای فهم اینکه آیا این قطعه شامل یک اتصال بلندتر از برخی طول‌های حداقل ثابت شده به نام یک seed هست یا نه، بررسی می‌شود (Dumas & Ninio, 1982).

آنالیز ژنوم‌های با تبار دور

محدودیت ابزارهای موجود در پردازش بازآرایی‌های ژنومی در مقیاس گسترده که عمدتاً در ژنوم‌های با فاصله تکاملی زیاد و حتی برخی ژنوم‌های نسبتاً نزدیک مشاهده می‌شود، چالش‌های تحلیلی متعددی ایجاد می‌کند. این نقیصه از یک سو توانایی شناسایی توالی‌های تشخیصی DNA که قادر به تشخیص تمام اعضای خانواده‌های ژنی بسیار متنوع باشند را محدود می‌سازد، و از سوی دیگر کارایی استفاده از هم‌ترازی‌های چندتایی را برای کشف مکانیسم‌های تکاملی که ممکن است دستخوش تغییرات معناداری شده باشند، کاهش می‌دهد. علاوه بر این، ابزارهای کنونی برای تشخیص شباهت‌ها در میان ژنوم‌های غیرمرتبط، نیازمند سطح بالایی از اشتراک توالی هستند و تنها در صورت تمرکز زیاد این تشابهات در مقابل گستره کامل ژنوم، قادر به انجام هم‌ترازی مؤثر خواهند بود (Yang et al., 2019). در جمع‌بندی، می‌توان ادعا کرد که تلفیق ابزارهای محاسباتی پیشرفته با پلتفرم‌های تصویرسازی یکپارچه، پارادایم جدیدی در مطالعات ژنومیک مقایسه‌ای ایجاد کرده که هم از نظر تحلیلی قدرتمند و هم از نظر کاربردی کارآمد است.

جمع‌بندی و چشم‌انداز

ژنومیک مقایسه‌ای علی‌رغم چالش‌های متعدد، به عنوان یک پارادایم اساسی در زیست‌شناسی مدرن به پیشرفت خود ادامه می‌دهد. آینده این حوزه در گرو توسعه الگوریتم‌های محاسباتی پیشرفته، یکپارچه‌سازی سامانه‌های اطلاعاتی و تکمیل بانک‌های داده ژنومی خواهد بود. با رفع این چالش‌ها، ژنومیک مقایسه‌ای پتانسیل آن را دارد که درک ما از مکانیسم‌های

گذار از نقشه‌های با وضوح پایین (مگابازی) به نقشه‌های با وضوح مولکولی (تک‌بازی) دانست.

مقایسه انتقادی یافته‌های اخیر نشان می‌دهد که حوزه ژنومیک مقایسه‌ای در حال گذار از یک رشته توصیفی به یک دانش پیش‌بین و کاربردی است. این گذار با سه تحول اصلی شامل تغییر مقیاس و حرکت از نگاه تک‌ژنومی محدود به دیدگاه جامع پان‌ژنومی، تغییر عمق و گذر از تحلیل‌های توالی‌محور صرف به یکپارچه‌سازی لایه‌های مختلف داده‌های آمیکی برای درک مکانیسمی و تغییر روش و جایگزینی تدریجی روش‌های سنتی با الگوریتم‌های هوش مصنوعی برای تفسیر داده‌های پیچیده همراه است. برای مطالعات آینده، ترکیب این سه رویکرد یعنی تهیه پان‌ژنوم‌های با کیفیت، ادغام داده‌های چند آمیکی و به‌کارگیری ابزارهای هوش مصنوعی ضروری است تا نه تنها الگوهای تکاملی را با وضوح بی‌سابقه‌ای آشکار کنیم، بلکه عملکرد زیستی آن‌ها را نیز به‌طور کمی پیش‌بینی و تأیید نماییم.

چارچوب تکاملی، شناسایی عناصر عملکردی: با شناسایی توالی‌های حفاظت‌شده و مناطق هم‌خطی، انتقال دانش بین گونه‌ها که امکان بهره‌گیری از داده‌های گونه‌های مدل برای درک زیست‌شناسی گونه‌های کمتر مطالعه‌شده را فراهم می‌سازد و یکپارچه‌سازی داده‌های چندسطحی یعنی ادغام داده‌های نقشه‌های ژنتیکی، فیزیکی و مولکولی، برای ایجاد درک جامعی از ساختار و عملکرد ژنوم دانست. محدودیت‌های ذاتی و چالش‌های پیش رو شامل نیاز به توسعه مدل‌های ریاضی و تکاملی پیشرفته که بتوانند پیچیدگی‌های تکاملی از جمله انتقال ژن جانبی را به‌طور دقیق در نظر بگیرند، یکپارچه‌سازی و تصویرسازی داده‌ها برای مدیریت و نمایش مؤثر حجم عظیم داده‌های چندبعدی حاصل از مطالعات ژنومیک، پر کردن شکاف‌های ژنومی یعنی تکمیل توالی‌یابی و تفسیر ژنوم گونه‌های کلیدی که حلقه‌های مفقوده درخت تکاملی محسوب می‌شوند و افزایش وضوح نقشه‌ها به معنی

References

- Alföldi, J. and Lindblad-Toh, K. 2013. Comparative genomics as a tool to understand evolution and disease. *Genome Research*, 23: 1063-1068. DOI: [10.1101/gr.157503.113](https://doi.org/10.1101/gr.157503.113)
- Alm, R., Ling, L. S., Moir, D. King, B. L., Brown, E. D., Doig, P. C., Smith, D. R., Noonan, B., Guild, B. C., DeJonge, B. L., Carmel, G., Tummino, P. J., Caruso, A., Uria-Nickelsen, M., Mills, D. M., Ives, C., Gibson, R., Merberg, D., Mills, S. D., Jiang, Q., Taylor, D. E., Vovis, G. F. and Trust, T. J. S. 1999. Genomic sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*, *Nature*, 397: 176-180. DOI: [10.1038/16495](https://doi.org/10.1038/16495)
- Anderson, L., Georges, M. 2004. Domestic-animal genomics: deciphering the genetics of complex traits, *Nature Reviews Genetics*, 5: 202-212. DOI: [10.1038/nrg1294](https://doi.org/10.1038/nrg1294)
- Argos, P., Kamer, G., Nicklin, M. J. and Wimmer, E. 1984. Similarity in gene organization and homology between proteins of animal picornaviruses and a plant comovirus suggest common ancestry of these virus families. *Nucleic Acids Research*, 1 (18): 7251-7267. DOI: [10.1093/nar/12.18.7251](https://doi.org/10.1093/nar/12.18.7251)
- Avsec, Ž., Agarwal, V., Visentin, D., Ledsam, J. R., Grabska-Barwinska, A., Taylor, K. R., ... and Kelley, D. R. 2021. Effective gene expression prediction from sequence by integrating long-range interactions. *Nature Methods*, 18(10): 1196-1203.
- Barcia-cruz, R., Balboa, S., Lema, A. and Romalde, J. 2024. Comparative genomics of *Vibrio toranzoniae* strains. *Research Square*, pp. 1-17. DOI: [10.1007/s10123-024-00557-z](https://doi.org/10.1007/s10123-024-00557-z)
- Chain, P., Kurtz, S., Ohlebusch, E. and Slezak, T. 2003. An applications-focused review of comparative genomics tools: capabilities, limitations and future challenges. *Briefing in bioinformatics*, 4: 105-123. DOI: [10.1093/bib/4.2.105](https://doi.org/10.1093/bib/4.2.105)
- Darling, A.E., Miklós, I., Ragan, M. A. 2008. Dynamics of Genome Rearrangement in Bacterial Populations, *PLOS Genetics*, pp. 4-7. DOI: [10.1371/journal.pgen.1000128](https://doi.org/10.1371/journal.pgen.1000128)
- Dumas, J. and Ninio, J. 1982. efficient algorithms for folding and comparing nucleic acid sequences, *Nucleic Acids Research*, 10: 197-206. DOI: [10.1093/nar/10.1.197](https://doi.org/10.1093/nar/10.1.197)
- Ellegren, H. 2008. Comparative genomics and the study of evolution by natural selection, *Molecular Ecology*, 17(21): 4586-4596. DOI: [10.1111/j.1365-294X.2008.03954.x](https://doi.org/10.1111/j.1365-294X.2008.03954.x)
- Flores, R., Huneau, C., Burlot, L., Lainé, M., Kimmel, E., Pommier, C., Alaux, M., Adam-Blondon, A. F., Pont, C., Quesneville, H. and Salse, J. 2023. SyntenyViewer: a comparative genomics-driven translational research tool. *Database: The Journal of Biological Databases and Curation*, 2023: baad027. DOI: [10.1093/database/baad027](https://doi.org/10.1093/database/baad027)
- Goffeau, A., Barrell, B., Bussey, G. H., Davis, R., Dujon, W. B., Feldmann, H., Galibert, F. J., Hoheisel, D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H. and Oliver, S. G. 1996. Life with 6000 genes, *Science*, 274(5287), 546: 563-567.
- Golicz, A. A., Bayer, P. E., Barker, G. C., Edger, P. P., Kim, H., Martinez, P. A., ... and Edwards, D. 2016. The pangenome of an agronomically important crop plant *Brassica oleracea*. *Nature Communications*, 7(1): 13390. DOI: [10.1038/ncomms13390](https://doi.org/10.1038/ncomms13390)
- Gottgens, B., Barton, L. M., Gilbert, J. G. Bench, A. J., Sanchez, M. J., Bahn, S., Mistry, S., Grafham, D., McMurray, A., Vaudin, M., Amaya, E., Bentley, D. R., Green, A. R. and Sinclair, A. M. 2000. Analysis of vertebrate SCL loci identifies conserved enhancers, *Nat, Biotechnology*, 18: 777-787. DOI: [10.1038/72635](https://doi.org/10.1038/72635)

- Gottgens, B., Gilbert, J., Barton, L. M., Grafham, L., Rogers, J. Bentley, D. R. and Green, A. R. 2001. Long-range comparison of human and mouse SCL loci: localized regions of sensitivity to restriction endonucleases correspond precisely with peaks of conserved noncoding sequences, *Genome Research*, 11: 87-97. DOI: [10.1101/gr.153001](https://doi.org/10.1101/gr.153001)
- Haft, D. H. 2015. Using comparative genomics to drive new discoveries in microbiology. *Current Opinion in Microbiology*, 23: 189-196. DOI: [10.1016/j.mib.2014.11.017](https://doi.org/10.1016/j.mib.2014.11.017)
- Hardison, R., Oeltjen, J. and Miller, W. 1997. Long human-mouse sequence alignments reveal novel regulatory elements: a reason to sequence the mouse genome, *Genome Research*, 7: 959-996. DOI: [10.1101/gr.7.10.959](https://doi.org/10.1101/gr.7.10.959)
- Hardison, R. C. 2003. Comparative genomics, *PLoS Biology*, 2: 52-58.
- Lund, J., Chen, F., Hua, Roe, B., Budarf, M., Emanuel, B. S. and Reeves, R. H. 2000. Comparative sequence analysis of 634kb of the mouse chromosome 16 region of conserved synteny with the human velocardiofacial syndrome region on chromosome 22q11.2, *Genomics*, 63: 374-383.
- Paulsen, I. T., Seshadri, R., Nelson, K. E., Eisen, J. A., Heidelberg, J. F., Read, T. D., Dodson, R. J., Umayam, L., Brinkac, L. M., Beanan, M. J., Daugherty, S. C., Deboy, R. T., Durkin, A. S., Kolonay, J. F., Madupu, R., Nelson, W. C., Ayodegi, B., Kraul, M., Shetty, J., Malek, J., Van Aken, S. E., Riedmuller, S., Tettelin, H., Gill, S. R. and White, O. 2000. The *Brucella suis* genome reveals fundamental similarities between animal and plant pathogens and symbionts, *Proceedings of the National Academy of Sciences, USA*, 20: 13148-13153. DOI: [10.1073/pnas.192319099](https://doi.org/10.1073/pnas.192319099)
- Read, T. D., Salzberg, S. L. and Pop, M. 2000. Comparative genome sequencing for discovery of novel polymorphisms in *Bacillus anthracis*, *Science*, 296: 2028-2033. DOI: [10.1126/science.1071837](https://doi.org/10.1126/science.1071837)
- Sengupta, S. and Azad, R. K. 2023. Leveraging comparative genomics to uncover alien genes in bacterial genomes. *Microbial Genomics*, 9(1): mgen000939. DOI: [10.1099/mgen.0.000939](https://doi.org/10.1099/mgen.0.000939)
- Safari, S. and Mehrabi, A. A. 2020. Association of microsatellite markers with root architecture and agromorphologic traits in diverse germplasm of bread wheat. *Plant Genetic Resources*, 18(6): 437-444.
- Tahmasebi, Z., Safari, S., Arminian, A., & Fatehi, F. 2022. Marker-trait association of Russian wheat aphid (*Diuraphis noxia*) resistance in the global diversity set of wild barely. *Journal of Plant Molecular Breeding*, 10(2): 46-60.
- Touchman, J. 2010. Comparative Genomics, *Nature Education Knowledge*, 3: 10-13.
- Varshney, R. K., Sinha, P., Singh, V. K., Kumar, A., Zhang, Q. and Bennetzen, J. L. 2020. 5Gs for crop genetic improvement. *Current Opinion in Plant Biology*, 56: 190-196.
- Wang, J., Chen, Y. and Zou, Q. 2023. Comparative Genomics and Functional Genomics Analysis in Plants *International Journal of Molecular Sciences*, 24: 6539. DOI: [10.3390/ijms24076539](https://doi.org/10.3390/ijms24076539)
- Wienberg, J., and Stanyon, R. 1998. Comparative chromosome painting of primate genomes, *ILAR*, J39: 77-91. DOI: [10.1093/ilar.39.2-3.77](https://doi.org/10.1093/ilar.39.2-3.77)
- Yang, X., Liu, D., Tschaplinski, T. J. and Tuskan, G. A. 2019. Comparative genomics can provide new insights into the evolutionary mechanisms and gene function in CAM plants. *Journal of Experimental Botany*, 70(22), 6539-6547. DOI: [10.1093/jxb/erz408](https://doi.org/10.1093/jxb/erz408)
- Yugi, K., Kubota, H., Hatano, A., and Kuroda, S. 2016. Trans-omics: how to reconstruct biochemical networks across multiple 'omic' layers. *Trends in Biotechnology*, 34(4): 276-290.
- Zhou, H., Clark, E., Guan, D., Lagarrigue, S., Fang, L., Cheng, H., Tuggle, C. K., Kapoor, M., Wang, Y., Giuffra, E., and Egidy, G. 2025. Comparative Genomics and Epigenomics of Transcriptional Regulation. *Annual Review of Animal Biosciences*, 13(1): 73-98. DOI: [10.1146/annurev-animal-111523-102217](https://doi.org/10.1146/annurev-animal-111523-102217)