ژنتیک نوین
دوره شانزدهم، شماره ٤، زمستان ١٤٠٠
صفحه ٤١٨ – ٤٠١

«مقاله مروری»

# تشخیص ویروس‌ها و ویروئیدهای بیماری‌زای گیاهی به‌کمک فناوری توالی‌یابی با توان بالا

# Detection of plant pathogenic viruses and viroids using high throughput sequencing technology

ارغوان علی سلطانی[١]، فریده فرح بخش*[٢،٣]، رضا لطفی[٤]، مجید طالبی[٥]

١- دانش‌آموخته پسا دکتری، گروه بیوتکنولوژی گیاهی، گروه، ویروس‌شناسی پزشکی، دانشکده پزشکی، دانشگاه کیپ تاون، کیپ تاون، آفریقای جنوبی

٢- دانش‌آموخته دکتری، گروه بیماری‌شناسی گیاهی، دانشکده کشاورزی، دانشگاه صنعتی اصفهان، اصفهان، ایران

٣- استادیار، گروه آفات و بیماری‌های گیاهی، مرکز تحقیقات کشاورزی و منابع طبیعی فارس، داراب، ایران

٤- دانشجوی دکتری، گروه بیوتکنولوژی، دانشکده کشاورزی، دانشگاه تربیت مدرس

٥- دانشیار، گروه بیوتکنولوژی، دانشکده کشاورزی، دانشگاه صنعتی اصفهان، اصفهان، ایران

**Alisoltani A[1], Farahbakhsh F[*2,3], Lotfi R[4], Talebi M[5]**

1- Postdoctoral Graduate, Department of Medical Virology, Faculty of Health Sciences, The University of Cape Town, Cape Town, South Africa
2- PhD Graduate, Department of Plant Pathology, Faculty of Agriculture, Isfahan University of Technology, Isfahan, Iran
3- Assistant Professor, Department of Plant Pests and Diseases, Agricultural Research, Education and Extension Organization (AREEO), Darab, Fars, Iran
4- PhD Student, Department of Biotechnology, Faculty of Agriculture, Tarbiat-Modares University, Tehran, Iran
5- Associate Professor, Department of Biotechnology, Faculty of Agriculture, Isfahan University of Technology, Isfahan, Iran

* نویسنده مسئول مکاتبات، پست الکترونیکی: farideh.farahbakhsh63@gmail.com

(تاریخ دریافت: ١٣٩٨/٠٧/١٥ – تاریخ پذیرش: ١٤٠٠/٠٧/٢٨)

## چکیده

**ویروس‌های گیاهی باعث بسیاری از بیماری‌های مهم می‌شوند و مسئول کاهش عملکرد و کیفیت محصول در سراسر جهان هستند. شناسایی سریع و دقیق ویروس‌ها در یک نمونه برای سرویس‌های بازرسی، قرنطینه و مدیریت بیماری مطلوب است. به‌تازگی، تکنیک‌های تعیین توالی نسل بعد (NGS) که نمایانگر یک رویکرد انعطاف‌پذیر در ردیابی پاتوژن گیاهی است، توسعه یافته است. از سال ٢٠٠٩، تکنیک‌های NGS در برخی نواحی ویروس شناسی گیاهی متشکل از توالی‌یابی ژنوم ویروس یا ویروئید، ردیابی و شناسایی، اکولوژی و اپیدمیولوژی آغاز به‌کار کرده است. تکنیک‌های NGS سریع، بسیار کارآمد و کم هزینه هستند و امکان توالی‌یابی چندین گیگابایت داده را در چند ساعت فراهم می‌سازند. در واقع، یک رویکرد عمومی برای شناسایی ویروس یا ویروئید است که نیازی به شناخت قبلی از میزبان یا ویروس ندارد. تنها با تعیین توالی کل آر.ان.ای‌های کوچک در یک اجرای NGS، شناسایی ویروس‌های دارای دی.ان.ای یا آر.ان.ای و ویروئیدها امکان‌پذیر می‌شود. در این مقاله، به پیشرفت‌های اخیر در به‌کارگیری تکنیک‌های NGS برای تشخیص ویروس‌ها و ویروئیدها در میزبانان آن‌ها پرداخته شده است.**

**واژه‌های کلیدی**

آسیب‌شناسی گیاهی
پلت فرم
توالی‌یابی نسل بعد شناسایی
ویروس
ویروئید

[ Downloaded from mg.genetics.ir on 2025-07-18 ]

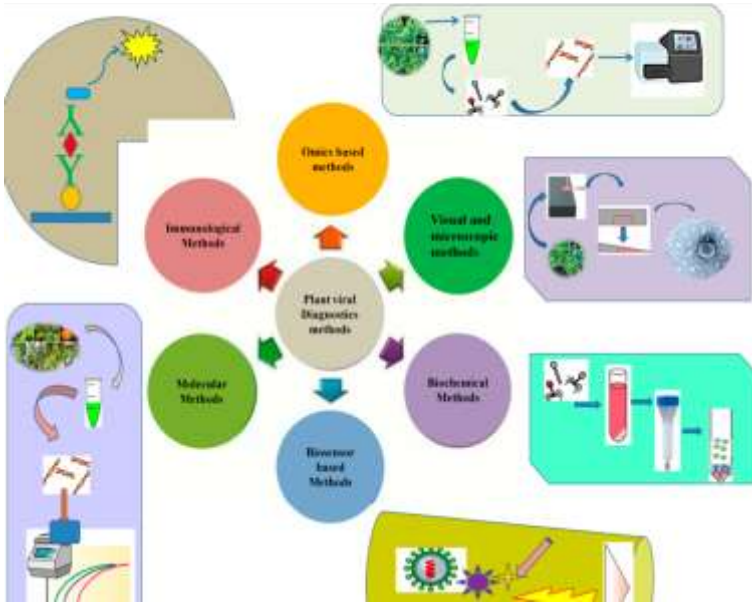
[ DOR: 20.1001.1.20084439.1400.16.4.1.5 ]

## مقدمه

ویروس‌ها بیمارگرهایی هستند که توانایی آلوده‌کردن تمامی موجودات را دارند. پیکره‌های ویروسی (ویریون‌ها[1]) ۴۵۰-۳۰ نانومتر اندازه دارند و ژنوم آن‌ها از نوع دی.ان.ای[2] و یا آر.ان.ای[3] می‌باشد که به‌وسیله پروتئین‌های کپسید[4] پوشیده شده و به‌صورت غلاف‌دار[5] و یا بدون غلاف[6] بسته‌بندی می‌شوند. ویروئیدها[7] نیز مولکول‌های کوچک حلقه‌ای از جنس آر.ان.ای و دارای ساختار ثانویه قوی هستند که باعث ایجاد بیماری‌های مهمی در گیاهان می‌شوند. تاکنون از روش‌های مختلفی از قبیل واکنش ایمنی‌سنجی آنزیمی یا الیزا[8]، واکنش زنجیره‌ای پلیمراز[9] و تکنیک‌های هیبریداسیون[10] نوکلئیک اسید مانند ریزآرایه‌ها[11]، به‌طور گسترده‌ای برای تشخیص سریع و ارزان ویروس‌ها و ویروئیدهای شناخته شده، در پزشکی و کشاورزی استفاده شده است (Delwart 2007; Mokili et al. 2012) (شکل ۱). اگرچه به‌دلیل اینکه این قبیل روش‌های تشخیصی، نیاز به معرف‌هایی[12] مانند آنتی بادی‌ها، آغازگرها[13] یا نشانگرها[14] دارند که فقط برای ویروس‌ها و ویروئیدهای شناخته شده موجود هستند، زمانی‌که بیماری توسط بیمارگری جدید یا مخلوطی از بیمارگرهایی که دارای تشابه توالی اندک هستند، ایجاد می‌شود، چنین تکنیک‌هایی بی‌اثر می‌باشند (Mokili et al. 2012).

تعیین توالی دی.ان.ای یکی از مهم‌ترین دستاوردهای بشر می‌باشد. طی دو دهه گذشته پیشرفت‌های زیادی در تکنولوژی‌های توالی‌یابی دی.ان.ای صورت گرفته که منجر به توسعه رهیافت‌های[15] جدید برای شناسایی ویروس‌ها و ویروئیدها شده است. متاژنومیکس[16] یکی از این راهکارهای نوین در میکروب‌شناسی محسوب می‌شود (Mokili et al. 2012). در تعریف ساده، متاژنومیکس، آنالیز جمعیت‌های میکروبی در نمونه‌های محیطی از طریق توالی‌یابی است (Roossinck et al. 2015).

[1] Virions
[2] Deoxyribonucleotide=DNA
[3] Ribonucleotides=RNA
[4] Capsid
[5] Enveloped
[6] non-enveloped
[7] Viroids
[8] Enzyme-linked immunosorbent assay =ELISA
[9] Polymerase chain reaction= PCR
[10] Hybridization
[11] microarray
[12] reagents
[13] Primers
[14] Probes
[15] approach
[16] Metagenomics




شکل ۱- روش‌های تشخیص رایج مورد استفاده برای شناسایی ویروس‌های گیاهی و تشخیص بیماری‌ها در گیاهان زراعی



[ Downloaded from mg.genetics.ir on 2025-07-18 ]
[ DOR: 20.1001.1.20084439.1400.16.4.1.5 ]

این رهیافت علاوه بر بهره‌مندی از روش‌های رایج تعیین توالی نوکلئیک اسید کل در نمونه‌های بیمار، از تکنولوژی نوین توالی‌یابی نسل بعد[1] استفاده می‌کند که پس از انجام توالی‌یابی بیمارگرها را با استفاده از ابزارهای بیوانفورماتیکی، تشخیص می‌دهد. به‌کمک تکنولوژی NGS، متاژنومیکس قابلیت شناسایی ویروس‌ها و ویروئیدهای جدید و از قبل شناخته شده در نمونه‌های بیمار را دارد و نیازی به اطلاعات اولیه در مورد بیمارگرها ندارد (Wu et al. 2015). به‌عنوان مثال تکنولوژی NGS، نقش کلیدی در شناسایی ویروس فلج حاد اسرائیلی[2] مربوط به عارضه فروپاشی کلونی در زنبور عسل داشته است (Cox-Foster et al. 2007). همچنین در علم پزشکی، با افزایش رو به رشد در مطالعات NGS، منبع جدید اطلاعات در کشف ژنومیکس کارکردی، بر اساس مارکرها فراهم شده است. تکنولوژی‌های RNA-seq و small RNA-seq، روش‌هایی نوین در کشف SSRهای بیان شده متمایز هستند که به‌عنوان مارکرهای زیستی در بیماری‌های مختلف کاربرد دارند. به‌عنوان مثال، با استفاده از SSR RNA-seq، مارکرهای زیستی جدیدی برای تشخیص سرطان پانکراس یافت شده است (Alisoltani et al. 2015b). پیشرفت‌های سریع در تکنولوژی NGS، هزینه و زمان شناسایی بیمارگر از طریق مسیر متاژنومیکس را به میزان زیادی کاهش داده و در سال‌های اخیر منجر به پیشرفتی شگرف در مطالعات متاژنومیکس ویروس‌ها و ویروئیدها در گیاهان شده است (Roossinck 2012). علاوه بر کاربرد RNA-seq در شناسایی ویروس‌ها و ویروئیدهای گیاهی که در قسمت‌های بعد این مقاله، به تفصیل شرح داده می‌شود، از رهیافت RNA-seq، به‌طور گسترده برای مطالعه پاسخ‌های ترانسکریپتوم گیاهی در تنش‌های محیطی مختلف استفاده شده است. مجموعه داده‌های مختلف RNA-seq، منابع باارزشی در توسعه مارکرهای SSR و انواع دیگر مارکرها در گونه‌های گیاهی هستند. به‌عنوان نمونه، اولین همگذاری مجدد[3] ترانسکریپتوم بادام در پاسخ به تنش سرما صورت گرفت و منجر به شناسایی هزاران ژن با بیان متمایز (differential expressed) شد. تعداد ۱۱ مارکر SSR RNA-seq به‌عنوان مارکرهای پلی مورفیک تشخیص داده شدند. همچنین تعدادی جایگاه[4] SSR بر روی ژن‌ها مشخص شدند که به تنش‌های سرما و سایر تنش‌های غیر زیستی واکنش نشان می‌دادند و عمدتاً شامل کالمودولین و فاکتور نسخه برداری تری هالوز GT-1-like بودند (Alisoltani et al. 2015a). در این مقاله به معرفی متاژنومیکس و شرح کاربردهای اخیر روش‌های متاژنومیکس وابسته به همولوژی و نیز مستقل از همولوژی در شناسایی و ردیابی ویروس‌ها و ویروئیدهای بیماریزای گیاهی در گیاهان پرداخته شده است.

متاژنومیکس که اولین بار توسط Handelsman et al. (1998) مطرح شد، کاربرد تکنیک‌های ژنتیکی جدید برای مطالعه جمعیت‌های میکروبی به‌طور مستقیم از یک محیط معین (خاک، اقیانوس، روده انسان و غیره)، بدون استفاده از کشت میکروبی و ایزوله کردن آن می‌باشد (Förstner 2009;Roossinck 2015; Steele And Streit 2006). این اصطلاح از مفهوم آماری meta-analysis مشتق شده است که عبارت است از فرایندی که در آن آنالیزهای آماری مجزا با هم ترکیب می‌شوند (Schloss 2003). در سال ۱۹۹۸ Handelsman و همکاران برای بررسی جمعیت کلی میکروب‌ها در خاک بدون کشت آن‌ها، اقدام به استخراج مستقیم دی.ان.ای از خاک نمودند، سپس دی.ان.ای را با آنزیم‌های محدود کننده[5] برش دادند و در داخل کروموزوم‌های مصنوعی باکتریایی (BAC) همسانه‌سازی[6] کردند. به‌علت اینکه این ناقل[7] قابلیت حمل قطعات بزرگ را در E.coli داشته، در ادامه آن‌ها کلون‌های BAC را برای بررسی فعالیت‌های بیولوژیکی و تولیدات طبیعی جدید مورد غربال قرار دادند (Handelsman et al. 1998). اطلاعاتی که در کتابخانه‌های متاژنومیک نگهداری می‌شوند، برای تعیین تنوع جمعیت و فعالیت آن، بررسی حضور یک میکروارگانیسم خاص یا مسیرهای بیوسنتز جدید و همچنین جستجو برای عملکردهای آنزیمی و یا یک ژن منفرد استفاده می‌شوند (Steele and Streit 2006; Förstner 2009).

[1] Next generation sequencing=NGS
[2] Israeli acute paralysis virus= IAPV
[3] De novo assembling
[4] Locus
[5] Restriction enzymes
[6] clone
[7] Vector



[ Downloaded from mg.genetics.ir on 2025-07-18 ]
[ DOR: 20.1001.1.20084439.1400.16.4.1.5 ]

شاخه اکولوژی میکروبی با معرفی فیلوژنتیک و از طریق مطالعه تنوع بر اساس *16s rDNA/rRNA* دچار دگرگونی شگرفی شد (Woese 2006). کاربرد آغازگرهای عمومی برای PCR مستقیم ژن‌های متنوع 16s از دی.ان.ای به‌همراه همسانه‌سازی و توالی‌یابی، داده‌های بسیاری را تولید می‌کند. به‌رغم اینکه آنالیز ژن *16s rRNA* اطلاعات ارزشمندی را در رابطه با تنوع و تکامل جمعیت میکروبی فراهم می‌کند، این روش نمی‌تواند تنوع بسیار پیچیده پروکاریوت‌ها را پوشش دهد. از طرفی این تکنیک نمی‌تواند نقش میکروارگانیسم‌های مختلف در جمعیت طبیعی را آشکار نماید. انواع روش‌ها بر پایه استفاده از *16s rDNA* و یا ژن‌های هدفمند برای متاژنومیک وجود دارند که در نهایت بر پایه توالی‌یابی محصول PCR می‌باشد. این داده‌های توالی‌یابی شده در ادامه مورد مقایسه قرار می‌گیرند و بررسی‌هایی از جمله بررسی‌های فیلوژنتیک یا بررسی پروفایل عملکرد روی آن‌ها انجام می‌گیرد. برای نمونه، در آخرین مطالعات صورت گرفته، تنوع و پروفایل عملکردی جمعیت‌های باکتریایی با استفاده از متاژنومیک با توان بالای ژن *16S rRNA* در نمونه‌های جمع آوری شده از سطوح محیطی مختلف در بیمارستان‌های آفریقای جنوبی بررسی شده است (Shobo et al. 2020). توالی‌یابی NGS، نوع دوم بررسی‌ها یعنی توالی‌یابی کل ژنوم بدون در نظر گرفتن ژن خاصی را امکان‌پذیر کرده است. در هر دو مورد توالی‌های محصولات PCR و توالی‌های کل ژنوم‌ها با این تکنیک‌ها قابل بررسی می‌باشند (Klindworth et al. 2012).

تکنیک‌های توالی‌یابی نسل بعد با سرعت و هزینه کمتر نسبت به تکنیک‌های نسل اول توالی‌یابی، بررسی و مقایسه ژنتیکی جوامع پیچیده را میسر کرده است. با این روش می توان جوامع با تنوع بیشتر همچون میکروب‌های مرتبط با جانوران مانند شکمبه گاو را بررسی نمود. تنها چالش کنونی این تکنیک، وجود پیچیدگی بسیار در نمونه‌های متاژنومیک است که تجزیه و تحلیل این داده‌ها را با مشکل مواجه ساخته است. در جدول ۱ انواع پلت فرم‌ها[1] که می‌تواند در بررسی داده‌های متاژنومیکس مورد استفاده قرار گیرد، نشان داده شده است (Wu et al. 2015).

[1] Platforms

معنی اصطلاح توالی‌یابی با توان بالا[2] با پیشرفت تکنولوژی‌های توالی‌یابی تغییر نموده است. تکنیک توالی‌یابی سانگر[3] که اولین بار در سال ۱۹۷۷ میلادی توسط فرد سانگر[4] ابداع شد، فرآیند بسیار پرزحمتی بود که تنها می‌توانست چندصد نوکلئوتید را در یک زمان مشخص توالی‌یابی کند. در این روش از دی دئوکسی نوکلئوتیدها[5] برای پایان همانندسازی و مشخص کردن موقعیت هریک از چهار باز تشکیل دهنده مولکول دی.ان.ای استفاده می‌شود. همانندسازی به کمک یک آغازگر که مکمل ناحیه ′۳ مولکول تک رشته دی.ان.ای است، آغاز می‌گردد و پس از رسیدن به دی دئوکسی نوکلئوتید مربوط به هر یک از چهار باز خاتمه می‌یابد. بدین ترتیب مولکول‌هایی با طول‌های مختلف به‌دست می‌آیند که با الکتروفورز از یکدیگر جدا و از روی باندهای تشکیل شده در ژل الکتروفورز، موقعیت بازها در رشته DNA مشخص شده و فرآیند توالی‌یابی صورت می‌گیرد (Daneshpour et al. 2014). در این روش توالی‌یابی، از رادیوایزوتوپ‌ها و مواد شیمیایی سمی نیز استفاده می‌شود. در اوایل دهه ۱۹۹۰، ابداع روش الکتروفورز موئین و توسعه سیستم‌های شناسایی مناسب، منجر به تولید توالی‌یاب HTS موئینه ۹۶ کانالی شد. بدین ترتیب سیستم‌های توالی‌یابی مویین سنگر رایج شد. مثالی از سیستم‌های توالی‌یابی موئین سنگر مدرن، DNA analyser 3730xl از Applied Biosystems (ABI) می‌باشد که ۹۶ یا ۳۸٤ توالی به طول تقریبی ۱۰۰۰-۶۰۰ نوکلئوتید در هر بار اجرا[6] تولید می‌کند و توانایی تولید حداکثر ۱/۵ مگاباز توالی در هر روز را دارد (جدول ۱) (Kircher and Keslo 2010). اساس کار با این دستگاه‌ها به‌ترتیب شامل الکتروفورز موئین برای جداسازی بر اساس طول قطعه، تشخیص و ثبت رنگ فلورسانس و در نهایت تولید داده می‌باشد. در این تکنیک، تکثیر قطعه مورد نظر، خالص سازی و تعلیق دوباره در محلول بافر قبل از ورود به دستگاه به‌صورت جداگانه صورت می‌پذیرد. تعدادی نرم‌افزار تجاری و غیر تجاری قادر به پیرایش توالی‌های دی.ان.ای به‌صورت خودکار

[2] High-throuphput sequencing=HTS
[3] Sanger
[4] Fred Sanger
[5] Dideoxynucleotides
[6] Run



[ Downloaded from mg.genetics.ir on 2025-07-18 ]
[ DOR: 20.1001.1.20084439.1400.16.4.1.5 ]

می‌باشند. این برنامه‌ها به کیفیت هر نگاره[1] نمره داده و نگاره‌های کم کیفیت (که به‌طور معمول در انتهاهای توالی قرار دارد) را حذف می‌کند (Daneshpour et al. 2014).

پلت فرم‌های توالی‌یابی جدید که از سال ۲۰۰۵ تولید شده‌اند، توانایی توالی‌یابی نسخه‌های مولکول دی.ان.ای تکثیر شده با تکنیک "wash and scan" را دارا هستند (Kircher and Kelso 2010). پلت فرم‌های توالی‌یابی جدید نسبت به تکنولوژی‌های توالی‌یابی سانگر عملکرد بهتری دارند به گونه‌ای که میزان داده تولیدی آن‌ها در یک روز، ۱۰۰۰-۱۰۰ برابر بیشتر است. به‌همین علت برای داده‌های حاصل از این پلت فرم‌ها از عبارت توالی‌یابی نسل بعد (NGS) به‌جای HTS استفاده شده است. روش‌های NGS (جدول ۱) اساس بیوشیمیایی متفاوتی دارند به‌گونه‌ای که از نظر آماده سازی الگو، توالی‌یابی، عکسبرداری و تجزیه و تحلیل داده‌ها با هم متفاوتند (Metzker 2010). یکی از تکنیک‌های توالی‌یابی نسل بعد تکنولوژی پیروسکانسینگ 454 است که به‌عنوان روش تعیین توالی حرارتی[2] نیز نام‌گذاری شده است. این روش توسط شرکت Life science آمریکا پیشنهاد شده و بعدها توسط شرکت Roche سوئیس خریداری شد. در این روش، قطعات دی.ان.ای در داخل قطرات آب در یک محیط روغنی تکثیر می‌شوند که به این روش تکثیر قطعه، امولسیون PCR[3] گفته می‌شود. در هر کدام از این قطرات آب، یک الگوی دی.ان.ای قرار دارد که به یک‌دانه[4] حاوی آغازگر متصل می‌شود و در نهایت بر اثر تکثیر آن یک کلونی فشرده ایجاد می‌شود. دستگاه توالی‌یاب حاوی تعداد زیادی چاهک با حجم در حد پیکولیتر است که هر کدام از این چاهک‌ها حاوی یک دانه و آنزیم مخصوص تعیین توالی است. این سیستم از آنزیم لوسیفراز[5] و سولفوریلاز[6] که تولید نور می‌کنند، برای تشخیص هر نوکلئوتید که به رشته دی.ان.ای تازه تشکیل شده افزوده می‌شود، استفاده می‌کند. میانگین تعداد خوانش[7]های تولید شده در هر بار انجام واکنش توسط توالی‌یاب GS FLX+_که از پلت فرم 454 استفاده می‌کند بیش از یک میلیون خوانش به طول تقریبی ۷۰۰ باز است، به گونه‌ای که هر بار انجام واکنش با این دستگاه تقریبا ۲۳ ساعت طول می‌کشد (Kircher and Kelso 2010). از این تکنیک به‌عنوان نمونه برای ردیابی ویروس X وانیل (Vanilla virus X) و ویروس نهفته وانیل (Vanilla latent virus) در وانیل استفاده شده است (Grisoni et al. 2017).

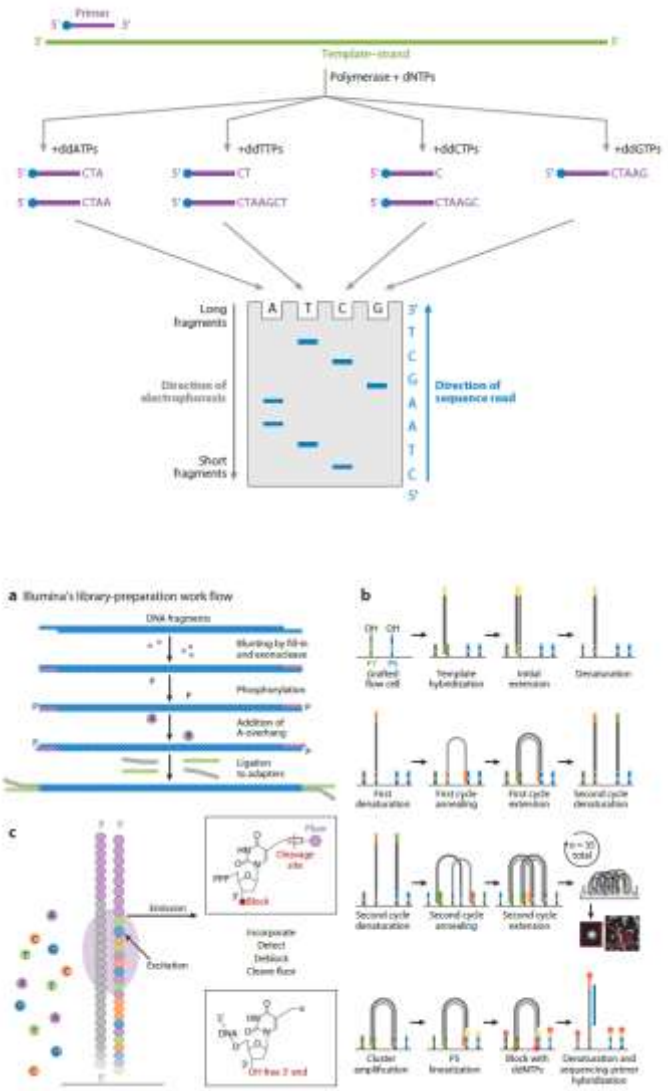
یکی دیگر از تکنیک‌های NGS، ایلومینا[8] نام دارد. این روش توسط شرکت آمریکایی Solexo ابداع شده و با توجه به اینکه این شرکت هم‌اکنون بخشی از شرکت Illimina آمریکا است به روش تعیین توالی ایلومینا[9] معروف شد. در این روش، استراتژی آماده سازی الگو و توالی‌یابی نسبتا با روش pyrosequencing 454 متفاوت است؛ به گونه‌ای که در آن از رنگ‌های پایان دهنده ولی برگشت‌پذیر[10] استفاده می‌شود. اساس کار بدین شکل است که ابتدا قطعات شکسته شده مولکول دی.ان.ای به آغازگرهای متصل به اسلاید وصل و به‌کمک تکنیک پل[11] تکثیر و در نهایت کلونی‌هایی متمرکز ایجاد می‌شوند. سپس چهار نوع باز متصل به رنگ‌های پایان دهنده ولی برگشت پذیر[12] به واکنش افزوده شده و هر باز در هر مرحله در صورتی که محل مناسبی برای اتصال وجود داشته باشد، به آن ناحیه متصل می‌شود. پس از اتصال تنها یک نوکلئوتید در هر مرحله به رشته الگو، نوکلئوتیدهای آزاد، شسته شده و از محیط واکنش حذف می‌شوند. بدین ترتیب فقط یک نوکلئوتید در هر مرحله از فرایند توالی‌یابی به دی.ان.ای افزوده می‌شود. در این لحظه دوربینی از نوکلئوتیدهای نشاندار شده با فلورسنت، عکس‌برداری می‌نماید. در پایان هر مرحله از فرایند خوانش، عامل متوقف‌کننده رشد رشته دی.ان.ای مکمل، توسط واکنش‌های شیمیایی از انتهای ′۳ جدا می‌شود و چرخه بعدی توالی‌یابی آغاز می‌شود (Metzker 2010) در شکل ۲، روش‌های توالی‌یابی سنگر و ایلومینا با هم مقایسه شده است.

[1] Peak
[2] Pyrosequencing
[3] Emulsion PCR
[4] Bead
[5] Luciferase
[6] Sulforylase
[7] Read

[8] Illumina
[9] Illumina(Solexa) Sequincing
[10] Reversible determinators
[11] Bridge Amplification
[12] RT-bases



[ Downloaded from mg.genetics.ir on 2025-07-18 ]
[ DOR: 20.1001.1.20084439.1400.16.4.1.5 ]




شکل ۲- مقایسه روش‌های توالی‌یابی سنگر و ایلومینا. توالی‌یابی سنگر (A)، توالی‌یابی ایلومینا (B).فرآیند ساخت کتابخانه ایلومینا (a)، تولید کلاستر ایلومینا از طریق تکثیر پل (b)، توالی‌یابی ایلومینا از طریق سنتز با پایانه‌های رنگ برگشت پذیر (c)

جدول ۱- مقایسه انواع پلت فرم‌های توالی‌یابی نسل بعد

| پلت فرم (Platform) | اساس | توان عملیاتی/واکنش (Throughput/run) | طول خوانش (Read length) | زمان/واکنش (Time/run) | نرخ هر بار اشتباه (Single-pass error rate) |
|---|---|---|---|---|---|
| ABI 3730xl | Sanger Sequencing | ٦٤ کیلوباز | ٩٠٠-٤٠٠ جفت باز | ١ ساعت | ٪ ٠/١-١ |
| GS FLX+ | SBS | ٧٠٠ مگاباز | ٧٠٠ جفت باز | ٢٣ ساعت | ٪١ |
| HiSeq 2500 | SBS | ١٠٠٠-٥ گیگاباز | ١٢٥×٢ جفت باز | ٦-١ روز | ٪ ٠/١ |
| MiSeq | SBS | ١٥-٠/٣ گیگاباز | ٣٠٠× ٢ جفت باز | ٥٥-٥ ساعت | ٪ ٠/١ |
| Ion Torrent | SBS | ٢-١/٢ گیگاباز | ٤٠٠ جفت باز | ٧/٣ ساعت | ٪١ |
| Ion Proton | SBS | ١٠ گیگاباز | ٢٠٠ جفت باز | ٤-٢ ساعت | ٪١ |
| Solid 5500XL | SBH | ٩٥ گیگاباز | ٦٠× ٢ جفت باز | ٦ روز | ٪٥ |
| PacBio RSII | SMRT | ١٠٠٠-٥٠٠ مگاباز | ١٥-١٠ کیلوباز | ٥-٠/٥ ساعت | ٪١٣ |

aSBS, sequencing by synthesis; SBH, sequencing by hybridization; SMRT, single molecule, real time



[ Downloaded from mg.genetics.ir on 2025-07-18 ]
[ DOR: 20.1001.1.20084439.1400.16.4.1.5 ]

برای مثال از پلت فرم Hiseq 2500 در ردیابی پنج ویروس انگور در آلودگی مخلوط گیاهان انگور از قبیل ویروس یک مرتبط با برگ قاشقی انگور (Grapevine leaf roll-associated virus 1) و ویروس سه مرتبط با برگ قاشقی انگور (Grapevine leaf roll-associated virus 3) (جنس آمپلوویروس[1]، خانواده کلستروویریده[2]) و سه ویروس خانواده Betaflexiviridae به نام‌های ویروس A انگور (*Grapevine virus* A)، ویروس B انگور (*Grapevine virus* B) و ویروس مرتبط با ساقه آبله‌ای rupestris انگور (Grapevine rupestris stem pitting associated virus) استفاده شده است (Xiao et al. 2019). همچنین از تکنیک Miseq برای ردیابی و تعیین توالی تعدادی ویروس‌ها از قبیل چهار ویروس مختلف در برگ‌های پاپایا مزارع کنیا شامل پوتی ویروس موزائیک هندوانه موروکو (Moroccan watermelon mosaic potyvirus) و کارلاویروس[3] ویروس پیسک خفیف نخود (Cowprea mild mottle virus)، و دو کارلاویروس فرضی دیگر مرتبط با ویروس رگبرگ روشنی خیار (Cucumber vein-clearing virus) به‌کار گرفته شده است (Mumo et al. 2020).

نوع دیگری از پلت فرم که در NGS استفاده می‌شود، پلت‌فرم ABI SOLiD نام دارد. این پلت‌فرم از محصولات شرکت Life Technologies محسوب می‌شود. در این پلت فرم، بر خلاف پلت فرم‌های ایلومینا و ٤٥٤ که در فرایند توالی‌یابی، واکنش رشد[4] قطعه را با آنزیم پلیمراز انجام می‌دهند، واکنش رشد قطعه به‌وسیله آنزیم‌های لیگاز انجام می‌شود. در این روش پیش از تعیین توالی، دی.ان.ای توسط دانه‌های پوشیده از آغازگر در حباب‌های روغنی (روش امولسیون PCR) تکثیر می‌شود. هر کدام از این حباب‌ها حاوی یک نسخه منحصر به‌فرد از مولکول دی.ان.ای است که پس از تکثیر در هر دانه، رشته‌های مشابه دی.ان.ای ایجاد می‌گردند و این دانه‌ها به‌منظور توالی‌یابی به سطح اسلایدهای شیشه‌ای ویژه‌ای متصل می‌شوند. در طی فرایند اتصال، الیگونوکلئوتیدهای دو رشته‌ای باز شده و در صورتی‌که جایگاه‌های مشابه خود را بیابند، به سطح اسلایدهای مزبور متصل می‌گردند. نتیجه تعیین توالی، کیفیت و طول قطعات در این روش، با روش ایلومینا قابل مقایسه است. اخیراً سیستم SOLiD 5500 xl، ۹۵ گیگا باز داده در هر واکنش توالی‌یابی طی ٦ روز با ترکیب خوانش‌هایی به‌طول ٦٠×٢ جفت‌باز تولید می‌کند (Kircher and Kelso 2010). به‌عنوان مثال از این تکنیک‌ها در ردیابی ویروس‌های مرکبات از قبیل تیپ ۲ سیتوپلاسمی ویروس لپروسیس مرکبات (Citrus leprosis virus sytoplasmic type 2) و ویروس توته‌ای رگبرگ مرکبات (Citrus vein enation virus) استفاده شده است (Loconsole et al. 2012).

تکنیک‌های توالی‌یابی جدید با عملکرد و کارایی بهبود یافته، در حال ایجاد و گسترش هستند (جدول ۱). اخیراً از اصطلاح توالی‌یابی نسل سوم نام برده می‌شود که شامل رهیافت‌هایی است که مولکول‌های دی.ان.ای بزرگ را بدون نیاز به وقفه بین مراحل خوانش، توالی‌یابی می‌کنند (Schadt et al. 2010). به‌طور کلی هدف اصلی نسل سوم تکنولوژی تعیین توالی، افزایش توان عملیاتی، کاهش هزینه‌ها، زمان و میزان مواد مصرفی طی فرایند توالی‌یابی است (Daneshpour et al. 2014). البته تکنیک‌هایی از قبیل Ion Torrent وجود دارند که گفته می‌شود بین نسل دوم و سوم توالی‌یابی قرار می‌گیرند، زیرا ویژگی‌های مربوط به هر دسته را به‌طور کامل دارا نیستند (Barzon et al. 2011). در سیستم Ion Torrent که فرایند توالی‌یابی در آن بر مبنای یک روش شیمیایی استاندارد است، از یک سیستم شناسایی نیمه رسانا استفاده می‌شود. این روش تعیین توالی بر خلاف روش نوری که در سایر سیستم‌ها استفاده می‌شود، از شناسایی یون‌های هیدروژن که در هنگام پلیمریزه شدن دی.ان.ای آزاد می‌شوند، استفاده می‌کند. چاهک‌های کوچک شامل الگوی دی.ان.ای هدف هستند که یک نوع از نوکلئوتیدهای آن آب‌گیری شده است. در شرایطی که نوکلئوتید معرفی شده به محیط، مکمل نوکلئوتیدهای دی.ان.ای هدف باشد، به رشته در حال رشد می‌پیوندد. این عمل باعث آزاد شدن یون‌های هیدروژن می‌شود. درنهایت اطلاعات حاصل از واکنش توالی‌یابی هر نوکلئوتید به‌وسیله یک حسگر فوق حساس یونی جمع‌آوری و ثبت می‌شود. در صورت وجود چند نوکلئوتید مشابه پشت سر هم در رشته الگو، بسته به تعداد نوکلئوتیدهای حاضر، شدت پیام‌های ارسالی نیز بیشتر می‌شود. به‌طور کلی

[1] Ampelovirus
[2] Closteroviridae
[3] Carlavirus
[4] Extension



[ Downloaded from mg.genetics.ir on 2025-07-18 ]
[ DOR: 20.1001.1.20084439.1400.16.4.1.5 ]

می‌توان گفت که با توجه به پیشرفت‌های مزبور در زمینه تکنولوژی توالی‌یابی، فرصت بی‌نظیری برای کشف بیمارگرها در گیاهان از طریق متاژنومیکس ویروسی فراهم شده است. به‌عنوان نمونه، از این تکنیک در آنالیز تنوع ژنتیکی ,ویروس کوچک گیلاس ۱ (Little cherry virus 1) در گونه‌های گیلاس شیرین و آلبالو استفاده شده است (Katsiani et al. 2018).

همچنین یکی دیگر از تکنولوژی توالی‌یابی تک مولکول‌ها، پلت فرم نانومنفذ[1] تحت عنوان Oxford nanopare می‌باشد که اخیراً در حال توسعه و تجاری سازی است. در این تکنولوژی، تغییرات مشخص جریان القایی هنگام عبور بازها از یک نانومنفذ بیولوژیکی اندازه‌گیری می‌شود، به‌صورتی که به یک پروتئین حرکتی مولکولی متصل شده است. در حال حاضر، دستگاه PromethION، با حدود ۱٤٤۰۰۰ کانال نانومنفذ، بالاترین توان در مقایسه با سه ابزار دیگر موجود در این سیستم عامل دارد (Mushtaq et al. 2020). در مجموع این تکنولوژی، از عملکرد پایین تر و خطای بالاتری نسبت به PacBio برخوردار است و به‌طور قابل توجهی توان کمتری نسبت به ایلومینا دارد. اگر نیازی به مجموعه داده‌های بزرگ نباشد، پلت فرمی از قبیل MinION می‌تواند نتایج مقرون به‌صرفه و در زمان واقعی را ارائه دهد (Villamor et al. 2019). از گروه پلت فرم‌های پیش تجاری، دو پلت فرم توالی‌یابی نسل دوم را می‌توان نام برد (Levy and Myers, 2016). اولی GENIUS (از Genapsys: http://www.genapsys.com) و دومی GeneReader (از Qiagen: https://www.qiagen.com/us/shop/sequencing/qiagen-genereader-platform). پلت فرم GENIUS از یک تراشه نیمه هادی مشابه Ion Torrent استفاده می‌کند و یک ابزار ارزان قیمت می‌باشد، در حالی‌که قادر است دی ان ا در اندازه ژنوم انسانی ($۳\times۱۰^{۹}$ جفت‌باز) را با قیمت کمتر از ۱۰۰۰ دلار توالی‌یابی کند (Villamor et al. 2019; Mehetre et al. 2021). در حال حاضر از این تکنیک برای ردیابی یا آنالیز ژنومی ویروس‌هایی از قبیل ویروس رگه ذرت (Maize streak virus)، ویروس موزائیک زردی ذرت (Maize yellow mosaic virus) و توتی ویروس[2] ذرت یا ویروس آبله آلو (Plum pox virus) استفاده شده است (Boronzo Badial et al. 2018; Mehetra et al. 2021). شکل ۱ مراحل آنالیز دو نوع داده شامل داده‌های حاصل از توالی‌یابی کل ژنوم و همچنین توالی‌یابی‌های محصولات PCR را نشان می‌دهد، هر دو مورد با تکنیک NGS توالی‌یابی می‌شوند. سپس بررسی‌های تاکسونومی و بررسی‌های پروفایل عملکرد بر روی داده‌های به‌دست آمده انجام شده و در نهایت نتایج هر دو با هم مقایسه می‌گردند.

گام اولیه در یک مطالعه متاژنومیکس، ساختن کتابخانه‌های متاژنومیکس است. اولین مرحله برای ساختن کتابخانه‌های متاژنومیکس استخراج دی.ان.ای از نمونه‌های محیطی برای مثال نمونه خاک، اقیانوس، روده انسان و غیره است. مولکول‌های دی.ان.ای به‌وسیله آنزیم محدود کننده مورد هضم قرار گرفته و طی مرحله اتصال[3] به داخل ناقلین[4] وارد می‌شوند. سپس این ناقل‌ها به داخل میزبان انتقال داده می‌شوند (Förstner 2009; Handelsman 2004). مرحله بعد پس از ساخت کتابخانه‌های متاژنومیکس، عملیات توالی‌یابی است که همان‌طور که در قسمت‌های قبل نیز اشاره شد از دو طریق انجام می‌گیرد. مسیر اول، توالی‌یابی تمام همسانه‌ها و دیگری ردیابی و بررسی یک ژن خاص همچون rDNA است (Förstner 2009). توالی‌یابی ژن rRNA نه تنها می‌تواند منجر به شناسایی ترکیب و انواع گونه در نمونه مورد نظر شود بلکه برای پیش‌بینی عملکرد موجودات در یک نمونه خاص می‌تواند استفاده شود (Abia et al. 2018). برای مثال، در جدیدترین تحقیقات صورت گرفته، با استفاده از توالی‌یابی *16S rRNA* و متاپروتئومیکس[5]، رابطه بین عملکرد میکروبی دستگاه تناسلی زنان و التهاب[6] در ۱۱۳ زن جوان و نوجوان آفریقای جنوبی در معرض خطر بالای عفونت HIV بررسی شد (Alisoltani et al. 2020). انواع داده‌های حاصل از تکنولوژی توالی‌یابی نسل بعد، به‌صورت روز افزون در حال افزایش هستند و تجزیه و تحلیل داده‌های متاژنومیکس و تفسیر آن‌ها اهمیت فراوانی دارد، بنابراین مرحله بعد استفاده از ابزارهای

[1] nanopore
[2] totivirus
[3] Ligation
[4] Vectors
[5] metaproteomics
[6] inflammation



[ Downloaded from mg.genetics.ir on 2025-07-18 ]
[ DOR: 20.1001.1.20084439.1400.16.4.1.5 ]

بیوانفورماتیکی است. روش آنالیز داده‌های حاصل از توالی‌یابی ژن rRNA شامل بررسی کیفیت داده‌ها و ردیف‌یابی توالی‌ها بر علیه رفرنس ژنوم‌ها همچون GreenGene 13 و Silva و تعیین کمیت و کیفیت گونه می‌باشد. این داده‌ها همچنین می‌تواند برای تعیین عملکرد توسط نرم‌افزارهایی همچون PICRUST مورد بررسی قرار گیرد (Abia et al. 2018). در هر حال آنالیز داده‌های متاژنومیکس حاصل از توالی‌یابی کل DNA موجود در نمونه دارای پیچیدگی بیشتر می‌باشد. به‌منظور آنالیز داده‌ها، در مرحله اول ردیف خوانش‌های ژنوم‌های میکروبی به داخل کانتیگ‌های طولانی به‌منظور تولید یک درفت از توالی‌های ژنومی همگذاری می‌شوند و فاصله[1] بین این کانتیگ‌ها برای کامل شدن درفت ژنومی به‌هم نزدیک می‌شوند. سپس برای همگذاری کردن توالی‌ها و همچنین همگذاری داده‌های NGS می‌توان از نرم‌افزارهایی بر پایه Kmer DeBruijn graph traversal از قبیل CLC bio، Velvet و سایر روش‌های نامبرده در جدول ۲ استفاده نمود. اکثر این نرم‌افزارها نیاز مند کامپیوتر دارای حافظه[2] بالا می‌باشند و برخی از همگذاری‌ها ممکن است چندین روز طول بکشد. مورد دیگر، خوانش‌هایی هستند که می‌توانند مستقیماً (بدون ساخت کانتیگ) برای دسته‌بندی‌های تاکسونومیکی و همچنین عملکردی مورد بررسی قرار گیرند (شکل ۱). البته تعداد کمی از مجموعه داده‌ها به این طریق می‌توانند طبقه‌بندی شوند (Scholz et al. 2011). تفسیر[3] داده‌های متاژنومیک، از کاربردی‌ترین مراحل تجزیه و تحلیل داده‌ها محسوب می‌شود.

[1] gap

[2] RAM
[3] Annotation

جدول ۲- برخی از انواع نرم‌افزارها و ابزارهایی که در بررسی متاژنومیکس مورد استفاده قرار می‌گیرند. برای بررسی و مطالعه هر نرم‌افزار می‌توانید به رفرنس‌های ذکر شده مراجعه کنید (Scholz et al. 2011)

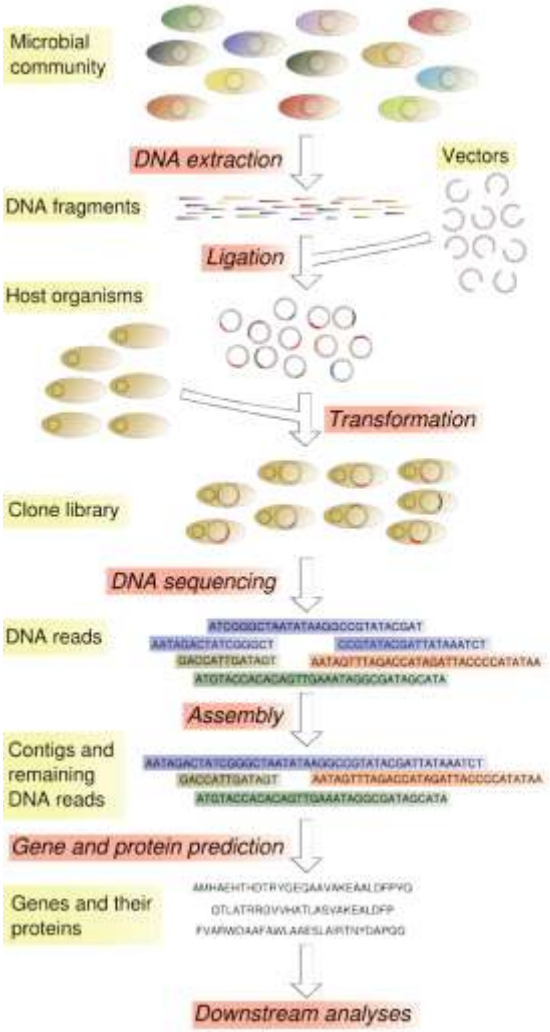
| | Software/Algorithm | Reference | Site link |
|---|---|---|---|
| Annotation and analysis | MG-RAST | (Aziz et al., 2008) | |
| | Gene-DB | (Logan-Klumpler et al., 2012) | |
| | IMG-M | (Markowitz et al., 2009) | |
| | Eragatis | (Orvis et al., 2010) | /http://ergatis.sourceforge.net |
| | DIYA | (Stewart et al., 2009)(28) | |
| | CloVR | - | http://clovr.org/ |
| | RATT | (Otto et al., 2011) | |
| | VMGAP | (Lorenzi et al., 2011) | |
| | CAMERA | (Seshadri et al., 2007)(31) | |
| | METAREP | (Goll et al., 2010) | |
| Assembly | RAY | (Boisvert et al., 2010) | |
| | Velvet | (Zerbino and Birney, 2008) | |
| | SOAPdenovo | (Chaisson and Pevzner, 2008) | |
| | Newbler | (Simpson et al., 2009) | |
| | ABySS | (Gnerre et al., 2011) | |
| | ALLPATHs | | |
| | Genovo | (Laserson et al., 2011) | |
| | CLCbio | | http://clcbio.com |
| | Meta-IDBA | (Peng et al., 2011) | |
| | MetaVelvet | | metavelvet.dna.bio.keio.ac.j |
| Mapping/alignment | BWA Bowtie | (Langmead et al., 2009) | |
| | SAMtools | (Li et al., 2009a) | |
| | SOAP2 | (Li et al., 2009b) | |
| | MrFAST | (Hach et al., 2010) | |
| | CloudBurst | (Schatz, 2009) | |
| | BFAST | (Homer et al., 2009) | |
| | MUMer | (Khan et al., 2009) | |
| | MOSAIK | | http://bioinformatics.bc.edu/marthlab/Mosaik |
| | BLAST | (Altschul et al., 1997) | |
| | HMM | (Eddy, 1998) | |
| Phylogenetic | PHYML | (Guindon and Gascuel, 2003) | http://www.lirmm.fr/w3ifa/MAAS/ |



[ Downloaded from mg.genetics.ir on 2025-07-18 ]
[ DOR: 20.1001.1.20084439.1400.16.4.1.5 ]

اگرچه تعیین تفسیر برای کانتیگ‌ها به‌علت گسترش تفسیر ژنوم‌های باکتریایی ساده به‌نظر می‌رسد، در حال حاضر هیچ سیستم مرکزی برای این منظور وجود ندارد. سیستم‌های مختلفی وجود دارد که در جدول ۲ به آن اشاره شده است. به‌عنوان مثال برای این کار می‌توان از سیستم مدیریتی گردش کار بر پایه وب به نام Ergatis استفاده نمود. این سیستم توسط Orvis و همکاران در سال ۲۰۱۰ معرفی شده است و کاربران را قادر می‌سازد که پایپ لاین[1] محاسباتی خود برای بررسی داده‌های ژنومیکس را نظارت کنند. این ابزار به‌صورت رایگان تحت آدرس وب http://ergatis.sourceforge.net/ قابل دسترسی است (Orvis et al. 2010). یکی از بارزترین مثال‌های کاربردی در زمینه متاژنومیکس، غربالگری آنزیمی در نمونه‌ها است. آنزیم‌های میکروبی به میزان زیادی در صنعت و زمینه‌های دیگر همچون تولید غذا استفاده می‌شوند. کتابخانه‌های دی.ان.ای ساخته شده از نمونه‌های محیطی، منابع مهمی برای یافتن انواع جدیدی از این آنزیم‌ها هستند (Guindon S and Gascuel 2003; Handelsman et al. 1998).

[1] pipeline




شکل ۳- مسیرهای مختلفی برای داده‌های متاژنومیکس وجود دارد و پایپ لاین مختلف توسط محققین تعریف شده‌اند. برای مثال در مسیری که در شکل نشان داده شده است، برای بررسی کتابخانه‌های پروتئینی از مدل مارکو (Eddy 1998) استفاده شده است و برای افزایش کارایی انتخاب غربال، درخت‌ها با بالاترین شباهت توسط روش (Guindon and Gascuel 2003) رسم شده‌اند. به کمک این پایپ لاین (Förstner 2009) توانست داده‌های متاژنومیک را مورد غربال قرار دهد و تعدادی آنزیم را به‌صورت موردی مطالعه کند (Förstner 2009).



[ Downloaded from mg.genetics.ir on 2025-07-18 ]
[ DOR: 20.1001.1.20084439.1400.16.4.1.5 ]

افزایش میزان داده‌های توالی‌یابی متاژنومیک، منبع اطلاعاتی جدیدی به‌منظور یافتن آنزیم‌های جدید با پتانسیل کاربردهای بیوتکنولوژی را فراهم می‌کند. آنزیم‌ها می‌توانند از طریق روش‌های غربال محاسباتی تشخیص داده شوند (Förstner 2009). مسیر کلی بررسی پروتئین بدین صورت است که پس از مرحله همگذاری[1] که در بالا توضیح داده شد، ژن‌ها و پروتئین‌ها تشخیص داده می‌شوند و سپس برای آنالیزهای دیگر مورد استفاده قرار می‌گیرند. شمای کلی در شکل ۳ نشان داده شده است. متاژنومیکس ویروسی شامل مطالعه ژنوم‌های ویروسی در نمونه‌های محیطی، گیاهی و جانوری است (Petrosino et al. 2009). از سال ۲۰۰۹، مطالعات مختلفی در رابطه با متاژنومیکس ویروس‌ها و ویروئیدهای گیاهی صورت گرفته است. به‌منظور شناسایی ویروس‌ها و ویروئیدها در یک مطالعه متاژنومیکس، یک سری مراحل کلی شامل آماده‌سازی نمونه، توالی‌یابی دی.ان.ای به‌کمک پلت فرم‌های NGS و در ادامه تجزیه و تحلیل‌های بیوانفورماتیک به‌منظور تفسیر نتایج توالی‌یابی وجود دارند (Xie et al. 2013). در متاژنومیکس ویروسی، روش‌های غنی‌سازی می‌توانند کارایی توالی‌یابی را به‌طور اختصاصی در ویروس‌ها و ویروئیدها افزایش دهند؛ به‌عبارت دیگر می‌توان با عملیات غنی‌سازی قبل از توالی‌یابی، نسبت مولکول‌های وابسته به ویروس‌ها و ویروئیدها را به سایر اسیدهای نوکلئیک غیر اختصاصی مرتبط با میزبان افزایش داد که در نهایت منجر به شناسایی بهتر و تعیین خصوصیت ویروس‌های جدید می‌شود. در ادامه پس از توضیح روش‌های غنی‌سازی توالی‌های ویروسی، به آنالیزهای بیوانفورماتیک بر روی توالی‌ها پس از عملیات توالی‌یابی، شناسایی وابسته به همولوژی و مستقل از همولوژی ویروس‌ها و ویروئیدها به کمک تکنولوژی‌های NGS و مشکلات موجود در زمینه متاژنومیکس در ویروس‌ها و ویروئیدها پرداخته می‌شود.

تاکنون شش ویروس گیاهی جدید با روش توالی‌یابی مستقیم آر.ان.ای کل از گیاهان میزبان شناسایی شده‌اند (جدول ۳). به‌منظور غنی‌سازی توالی‌هایی با دقت و عمق زیاد[2] توالی‌یابی و نیز بهبود فرایند ردیابی ویروس‌های جدید، چندین روش اختصاصی برای ویروس‌ها و ویروئیدها ایجاد و توسعه یافته‌اند. به‌عنوان مثال حذف مولکول‌های rRNA میزبان از آر.ان.ای کل، یک فرایند غنی‌سازی است که نسبت خوانش‌های مربوط به ویروس‌ها و ویروئیدها را طی یک فرایند توالی‌یابی، نسبت به خوانش‌های مرتبط با میزبان افزایش خواهد داد. استراتژی‌های غنی‌سازی منجر به کشف هفت جنس ویروس‌های آر.ان.ای دار جدید و دو بادناویروس[3] دارای ژنوم دی.ان.ای شده است. به‌صورت کلی می‌توان گفت که غنی‌سازی مولکول‌های ذکر شده در مطالعات متاژنومیکس ویروسی رایج‌تر است (Guindon S and Gascuel 2003;Gu et al. 2014).

از آنجاکه آر.ان.ای دو رشته‌ای (dsRNA) به‌عنوان واسطه همانندسازی توسط ویروس‌های آر.ان.ای دار و ویروئیدها سنتز شده و گیاهان به‌طور معمول dsRNA تولید نمی‌کنند، توالی‌یابی dsRNA کل می‌تواند به میزان زیادی نسبت خوانش‌های اختصاصی ویروس‌ها و ویروئیدها را افزایش دهد (Wylie et al. 2013). به‌عنوان نمونه، در مقایسه توالی‌یابی با دقت و عمق زیاد بین آر.ان.ای کل و dsRNA، پس از غنی‌سازی dsRNA از نمونه گیاهی، میزان خوانش‌های ویروسی از ۲٪ به ۵۳٪ افزایش یافت. به‌منظور خالص‌سازی dsRNA، ستون‌های جداکننده ویژه (به‌عنوان مثال CF-11 spin column) در دسترس هستند. اخیراً به‌منظور جداسازی dsRNA، سیستم کروماتوگرافی تبادل آنیونی با استفاده از ستون‌های مونولیتیک ایجاد شده است (Gu et al. 2014). شش ویروس از هفت ویروس کشف شده از طریق توالی‌یابی dsRNA، ژنوم آر.ان.ای دارند. تعداد کمتری از ویروس‌های دی.ان.ای دار گیاهی با این روش شناسایی شده‌اند؛ چراکه این ویروس‌ها در چرخه زندگی خود به مقدار کافی dsRNA طویل تولید نمی‌کنند (Al Rwahnih et al. 2013).

[1] Assembly
[2] Deep sequencing
[3] Badnavirus



[ Downloaded from mg.genetics.ir on 2025-07-18 ]
[ DOR: 20.1001.1.20084439.1400.16.4.1.5 ]

جدول۳-ویروس‌ها و ویروئیدهای گیاهی جدید کشف شده توسط روش‌های NGS

| ویروس های آر.ان.ا دار | خانواده(جنس) | عفونی |
|---|---|---|
| Gayfeather mild mottle virus (GMMV) | Bromoviridae (Cucumovirus) | + |
| Grapevine Syrah virus-1 (GSyV-1) | Tymoviridae (Marafivirus) | - |
| Cassava brown streak virus (CBSV) | Potyviridae (Ipomovirus) | - |
| Hardenbergia virus (HarVA) | Betaflexiviridae | – |
| Lettuce necrotic leaf curl virus (LNLCV) | Secoviridae (Torradovirus) | + |
| Yellow tailflower mild mottle virus (YTMMV) | Virgaviridae (Tobamovirus) | + |
| Caladenia virus A (CalVA) | Potyviridae (Poacevirus) | – |
| Donkey orchid virus A (DOVA) | Potyviridae (Potyvirus) | – |
| Apple rubbery wood virus 1 and 2 | Phenuiviridae | - |
| Blackberry leaf mottle-associated virus | Fimoviridae (Emaravirus) | |
| Diuris virus A (DiVA) | Betaflexiviridae (Capillovirus) | - |
| Diuris virus B (DiVB) | Betaflexiviridae (Capillovirus) | - |
| Diuris pendunculata cryptic virus (DPCV) | Partitiviridae | – |
| Ipomoea batatas Rhabdovirus N-like sequences (IbRNLS) | Rhabdoviridae | – |
| Cymbidium mosaic virus (CymMV) | Alphaflexiviridae (Potexvirus) | - |
| Raspberry latent virus (RpLV) | Reoviridae (Reovirus) | + |
| Grapevine virus F (GVF) | Betaflexiviridae (Vitivirus) | + |
| Donkey orchid symptomless virus (DOSV) | شناخته نشده | + |
| Persimmon virus A (PeVA) | Rhabdoviridae (Cytorhabdovirus); unassigned | + |
| Persimmon latent virus (PeLV) | Rhabdoviridae (Cytorhabdovirus); unassigned | + |
| Eggplant mild leaf mottle virus (EMLMV) | Potyviridae | + |
| Pepper yellow leaf curl virus (PYLCV) | Luteoviridae (Polerovirus) | + |
| Tomato necrotic stunt virus (ToNSV) | Potyviridae (Potyvirus) | + |
| Grapevine pinot gris virus (GPGV) | Betaflexiviridae (Trichovirus) | – |
| Citrus yellow vein clearing virus Y1 (CYVCV-Y1) | Alphaflexiviridae (Mandarivirus) | + |
| Yam bean mosaic virus (YBMV) | Potyviridae (Potyvirus) | + |
| Woolly burdock yellow vein virus (WBYVV) | شناخته نشده | – |
| sweet potato C6 virus (SPC6V) | Betaflexiviridae (Carlavirus) | + |
| Andean potato mild mosaic virus (APMMV) | Tymoviridae (Tymovirus) | + |
| Tomato mottle mosaic virus (ToMMV) | Virgaviridae (Tobamovirus) | + |
| Citrus leprosis virus cytoplasmic type 2 (CiLV-C2) | Cilevirus | + |
| Citrus vein enation virus (CVEV) | Luteoviridae (Enamovirus) | + |
| Tomato matilda virus (TMaV) | Iflaviridae (Tomavirus) | + |
| Cassava polero-like virus (CsPLV) | Luteoviridae (Polerovirus); | + |
| Cassava new alphaflexivirus (CsNAV) | Alphaflexiviridae(Potexvirus) | + |
| Cassava torrado-like virus (CsTLV) | Secoviridae (Torradovirus) | + |
| Rose leaf rosette-associated virus (RLRaV) | Closteroviridae (Closterovirus) | - |
| ویروس های دی.ان.ا دار | | |
| Sweet potato badnavirus C1 (SPBV-C1) | Caulimoviridae (Badnavirus); | - |
| Sweet potato badnavirus C2 (SPBV-C2) | Caulimoviridae (Badnavirus) | - |
| Grapevine red blotch-associated virus (GRBaV) | Geminiviridae | |
| Sugarcane white streak virus (SWSV) | Geminiviridae (Mastrevirus) | + |
| Piper yellow mottle virus (PYMoV) | Caulimoviridae (Badnavirus) | + |
| Piper DNA virus 1 (PDV-1) | Caulimoviridae | - |
| Piper DNA virus 2 (PDV-2) | Caulimoviridae | - |
| ویروس های آر.ان.ا دار | | |
| Two badnaviruses | | - |
| One mastrevirus | | |
| Grapevine vein clearing virus (GVCV) | | + |
| Citrus chlorotic dwarf-associated virus (CCDaV) | | + |
| Grapevine geminivirus (GVGV) | | – |
| Pagoda yellow mosaic associated virus (PYMAV) | | – |
| ویروئیدها | | |
| Persimmon viroid 2 (PVd2) | | + |
| Grapevine hammerhead viroidlike RNA (GHVd | | - |
| Apple hammerhead viroid-like RNA (AHVd-like | | + |
| Grapevine latent viroid (GLVd) | | - |

[ Downloaded from mg.genetics.ir on 2025-07-18 ]

[ DOR: 20.1001.1.20084439.1400.16.4.1.5 ]

ژنوم ویروس به‌صورت آر.ان.ای یا دی.ان.ای در پیکره‌های شبه ویروسی[1] بسته‌بندی شده تا از آسیب آنزیم‌های DNase و RNase حفاظت شود. رهیافت غنی‌سازی پیکره‌های شبه ویروسی با روش‌هایی نظیر هموژنیزه کردن، فیلتراسیون و اولتراسانتریفوژ، به‌طور وسیعی برای کشف ویروس‌ها در نمونه‌های محیط‌زیستی، اقیانوسی، مدفوعی و گیاهی استفاده شده است (Delwart 2007; Mokili et al. 2012). اینوکلوم پیکره‌های شبه ویروسی دارای آلودگی‌های باکتریایی و میتوکندریایی هستند؛ بنابراین اغلب تحت تیمار با کلروفرم قرار می‌گیرند تا غشاهای باکتریایی و میتوکندریایی قبل از استخراج نوکلئیک اسیدهای مرتبط با پیکره‌های شبه ویروسی، به‌منظور انجام توالی‌یابی با دقت و عمق بالا، تخریب گردند. متاسفانه ویروس‌های غلاف دار حساس به تیمار کلروفرم هستند. علاوه بر این چون خالص‌سازی موفق ویریون تعدادی از ویروس‌ها، نیازمند ایجاد و توسعه روش‌های اختصاصی است، دستیابی به یک روش یکسان غنی‌سازی پیکره‌های شبه ویروسی برای همه ویروس‌ها ممکن نیست. با این وجود، توالی‌یابی با عمق و دقت زیاد نوکلئیک اسیدهای مشتق شده از پیکره‌های شبه ویروسی استخراج شده از گیاهان، منجر به کشف چهار ویروس دی.ان.ای دار شامل ویروس مصری رگه‌ای نیشکر ((Sugarcane streak Egypt virus(SSEV))، ویروس پیسک زردی فلفل سیاه ((Piper yellow mottle virus(PYMV))، ویروس دی.ان.ا-۱ فلفل سیاه ((Piper DNA virus-1(PDV-1)) و ویروس دی.ان.ای-۲ فلفل سیاه ((Piper DNA virus-2(PDV-2)) و دو ویروس آر.ان.ای دار به نام‌های ویروس پیسک برگ خفیف بادمجان((Eggplant mild leaf mottle virus (EMLMV)) و ویروس پیچیدگی برگ زرد فلفل ((Pepper yellow leaf curl virus (PYLCV)) شده است (Wu et al. 2015).

ویروس پیچیدگی برگ زرد فلفل ((Pepper yellow leaf curl virus (PYLCV)) شده است (Wu et al. 2015). به‌علت مقادیر بسیار پایین نوکلئیک اسید کل به‌دست آمده از اینوکلوم‌های پیکره‌های شبه ویروسی و dsRNA، تکثیر نوکلئیک اسیدهای استخراج شده از طریق واکنش PCR یا RT-PCR به روش مستقل از توالی[2]، قبل از ساختن کتابخانه‌های مورد نیاز برای توالی‌یابی با دقت و عمق زیاد، لازم است. برای تکثیر اسید نوکلئیک در ویروس‌هایی با ژنوم دی.ان.ای دو رشته‌ای (dsDNA) حلقوی، تکنیک دایره غلتان بهبود یافته مؤثر می‌باشد (Cox-Foster et al. 2007; Mokili et al. 2012). همچنین مشخص شده است که گیاهان و بی‌مهرگان در واکنش به آلودگی با ویروس‌ها (صرف نظر از این که ژنوم آر.ان.ای یا دی.ان.ای باشد)، مولکول‌های آر.ان.ای کوچک مداخله‌گر[3] مشتق شده ویروسی را تولید می‌کنند (Adams et al. 2013). در سلول‌های پستانداران نیز تولید فراوان مولکول‌های siRNA ویروسی پس از آلودگی به دو ویروس رشته‌ای آر.ان.ای مثبت غیرخویشاوند مشاهده شده است (Romanovskaya et al. 2013). علاوه بر این، واسطه‌های همانندسازی dsRNA ویروئیدها و مولکول‌های آر.ان.ای ماهواره‌ای[4]، در گیاهان به شکل مولکول‌های siRNA فراوری می‌شوند (Adams et al. 2013). اولین مطالعه توالی‌یابی با دقت و عمق زیاد بر روی مولکول‌های siRNA ویروسی نشان داد که مولکول‌های siRNA ویروسی به میزان زیادی با همدیگر همپوشانی دارند. این درحالی است که این مولکول‌ها فقط ۲۴-۲۱ نوکلئوتید طول دارند. تولید فراگیر siRNA ویروسی در میزبانان یوکاریوتی متنوع و مشخصه همپوشانی مولکول‌های siRNA ویروسی سبب ایجاد یک استراتژی جدید برای کشف ویروس‌ها از طریق غنی‌سازی و توالی‌یابی مولکول‌های آر.ان.ای کوچک[5] شده است. در این رهیافت که به‌عنوان کشف ویروس از طریق توالی‌یابی با دقت و عمق زیاد و چیدمان مجموع توالی مولکول‌های آر.ان.ای کوچک[6] (به اختصار vdSAR) شناخته می‌شود، مولکول‌های آر.ان.ای کوچک از بافت‌ها و سلول‌های بیمار برای توالی‌یابی با دقت و عمق زیاد از طریق پلت فرم‌های NGS، جدا شده و پس از توالی‌یابی به کانتیگ‌ها[7] و قطعات توالی

[1] Virus-like particles (VLPs)

[2] Sequence-independent
[3] Small interfering RNA= siRNA
[4] Satellite
[5] Small RNA
[6] virus discovery by deep sequencing and assembly of total host small RNAs
[7] Contigs



[ Downloaded from mg.genetics.ir on 2025-07-18 ]
[ DOR: 20.1001.1.20084439.1400.16.4.1.5 ]

بزرگ همگذاری می‌شوند. سپس از این کانتیگ‌ها برای کشف ویروس استفاده می‌شود (Ding and Lu 2013).

در پروتکل‌های رایج ساخت کتابخانه از مولکول‌های آر.ان.ای کوچک، به خالص‌سازی از ژل مولکول‌های مزبور، قبل یا پس از اتصال آداپتورهای ′۳ و ′۵ نیاز نیست و عملیات ساخت کتابخانه می‌تواند طی یک روز انجام گیرد (Romanovskaya et al. 2013). تکثیر مستقل از توالی مورد نیاز به‌منظور غنی‌سازی dsRNA و پیکره‌های شبه ویروسی، برای مولکول‌های آر.ان.ای کوچک ضروری نیست. بنابراین آماده‌سازی نمونه جهت توالی‌یابی مولکول‌های آر.ان.ای کوچک، نسبت به پروتکل‌های خالص‌سازی dsRNA و پیکره‌های شبه ویروسی به لحاظ تکنیکی آسان‌تر بوده و به زمان کمتری نیاز دارد (Romanovskaya et al. 2013). دلایل زیادی وجود دارند که باعث شده است تا محققین، غنی‌سازی و توالی‌یابی مولکول‌های آر.ان.ای کوچک را نسبت به سایر استراتژی‌های غنی‌سازی در جهت کشف ویروس‌ها و ویروئیدهای گیاهی ترجیح دهند. دلیل اول این که همانندسازی ویروس‌های دی.ان.ای دار و آر.ان.ای دار و نیز عوامل شبه ویروسی از قبیل ویروئیدها و مولکول‌های آر.ان.ای ماهواره ای در گیاهان، تجمع بسیار زیاد siRNAهای اختصاصی بیمارگر را القا می‌کند که تا ۳۰٪ مولکول‌های آر.ان.ای کوچک توالی‌یابی شده در گیاهان بیمار را شامل می‌شود (Adams et al. 2013). به‌علت اینکه میزان توالی‌یابی و پیچیدگی داده‌ها به میزان زیادی کاهش یافته، توالی‌یابی نمونه‌های چندتایی که به‌وسیله بارکدها در یک مسیر[1] انفرادی نشاندار شده‌اند، عمق کافی در شناسایی بیمارگر را فراهم می‌سازد. بنابراین کشف ویروس از طریق توالی‌یابی با عمق و دقت زیاد و همگذاری آر.ان.ای‌های کوچک کل، مقرون به‌صرفه و سودآور است. دلیل دوم، برخلاف اینوکلوم‌های پیکره‌های شبه ویروسی و dsRNA، تمام ویروس‌ها و ویروئیدهای در حال همانندسازی در گیاه بیمار، به‌کمک توالی‌یابی با عمق و دقت زیاد یک کتابخانه واحد از مولکول‌های آر.ان.ای کوچک تشخیص داده می‌شوند. این خصوصیت به‌منظور کاربردهای متعدد در تشخیص بیماری و برنامه‌های قرنطینه قابل توجه است. علت سوم این است که مولکول‌های siRNA ویروسی و شبه ویروسی، محصولات یک پاسخ ایمنی فعال میزبان نسبت به آلودگی هستند و به‌علت بیوژنز آن‌ها توسط پروتئین‌های دایسر اختصاصی، الگوهای اختصاصی توزیع اندازه[2] در گونه‌های میزبانی خاص نشان می‌دهند (Adams et al. 2013). بنابراین الگوی توزیع اندازه مولکول‌های آر.ان.ای کوچک می‌تواند نشان دهد که آیا بیمارگر ویروسی یا شبه ویروسی شناسایی شده، فعالانه در گیاه یا در گونه حشره‌ای یا قارچی میزبان همانندسازی و تکثیر می‌کند یا خیر (Li et al. 2013). سرانجام، داده کاوی[3] کتابخانه‌های مشابه مولکول‌های آر.ان.ای کوچک می‌تواند منجر به کشف بیمارگرهای نوظهور شود که هیچگونه تشابه توالی قابل تشخیصی به‌کمک ابزارهای انفورماتیکی (اطلاعاتی) قابل دسترس و رایج نشان نمی‌دهند (Zhang et al. 2014).

قبل از همگذاری توالی و شناسایی بیمارگر، داده‌های خام[4] تولید شده به‌وسیله پلت فرم‌های NGS باید پیش پردازش شوند تا آداپتورها و توالی‌های با کیفیت پایین حذف گردند (شکل۳). کنترل کیفیت به نوع تکنولوژی توالی‌یابی مورد استفاده بستگی دارد. پارامترهای استاندارد و آستانه‌ها معمولا توسط سازنده[5] فراهم می‌شوند. در حال حاضر فرایند استاندارد سازی بر مبنای تکنولوژی‌های قدیمی‌تر مانند 454/Roche pyrosequencing یا Illumina sequencing by synthesis از طریق سنتز وجود دارد. برای توالی‌یابی چندگانه کتابخانه‌های مخلوط در یک مسیر واحد، مرحله اضافه چندگانه‌زدایی[6] با استفاده از بارکدهای ساخته شده در آغازگرهای PCR، پیش از همگذاری توالی ضروری است. زمانی که توالی ژنوم گیاه میزبان در دسترس باشد، حذف توالی‌های اختصاصی میزبان قبل از همگذاری به‌کمک نرم‌افزار (شکل ٤)، تجزیه و تحلیل‌های بیوانفورماتیکی بعدی را تسریع می‌نماید (Mokili et al. 2012). عملیات حذف آداپتور و فیلتر کردن توالی‌های میزبان می‌تواند منجر به ایجاد مصنوعات[7] دیگر شود (Roossinck 2012).

[1] Lane

[2] Size distribution
[3] Data mining
[4] Raw
[5] manufacturer
[6] demultiplexing
[7] artifacts



[ Downloaded from mg.genetics.ir on 2025-07-18 ]
[ DOR: 20.1001.1.20084439.1400.16.4.1.5 ]

همگذاری توالی‌های پیش پردازش شده با چندین الگوریتم اصلی قابل اجرا است. این الگوریتم‌ها شامل Velvet، Oases و Vcake است که در دسترس عموم قرار دارند (Wu et al. 2010). پارامترهای مورد استفاده در همگذاری توالی، مشابه پارامترهای همگذاری ژنوم هستند و از طریق الگوریتم به‌کار گرفته شده، تعریف می‌گردند. پس از همگذاری، توالی‌های کانتیگ از طریق ابزارهای جستجوی تشابه[1] با توالی‌های از قبل ثبت شده در پایگاه‌های اطلاعاتی محلی یا پایگاه‌های اطلاعاتی عمومی از قبیل بانک ژن، مقایسه می‌شوند. یکی از روش‌های رایج مقایسه توالی‌های همگذاری شده با پایگاه اطلاعاتی نوکلئوتیدی بانک ژن، استفاده از بسته[2] BLAST است (Mokili et al. 2012). دو برنامه غالب مورد استفاده به‌منظور مقایسه در سطوح نوکلئوتیدی و آمینواسیدی به ترتیب BLASTn و BLASTx هستند. زمانی‌که کانتیگ‌ها، تشابه بالا (بیش از ۹۰٪ یکسانی[3] و ۸۵٪ پوشش[4]) با یک ویروس شناخته شده نشان می‌دهند، ویروس‌های جدید و شناخته شده به راحتی از یکدیگر تمیز داده می‌شوند (Ding and Lu 2011). در صورتی‌که یک کانتیگ، همولوژی اندکی با ویروس شناخته شده به‌خصوص در سطح پروتئین نشان دهد، کانتیگ اغلب نشان دهنده ویروس جدیدی است که از نظر تاکسونومیکی تنها در سطح یک خانواده ویروسی می تواند به رسمیت شناخته شود (Ding and Lu 2011).

یکی از مشکلات موجود در مسیر متاژنومیکس ویروسی، تفسیر مجدد[5] هزاران کانتیگ همگذاری شده است. زمان محاسبه مورد نیاز برای تفسیر، به موازات رشد تصاعدی پایگاه‌های اطلاعاتی توالی، افزایش می‌یابد. استفاده از پایگاه‌های اطلاعاتی اختصاصی یا زیرمجموعه‌های بانک ژن می‌تواند به حل این مشکل کمک کند. با توجه به تقاضای فراوان برای تفسیر مقادیر عظیم داده‌های به‌دست آمده از توالی‌یابی NGS، برنامه‌های کامپیوتری از قبیل USEARCH و HHbits که به مراتب سریع‌تر از BLAST هستند، ایجاد و توسعه یافته‌اند (Wu et al. 2010).

[1] Homology
[2] Package
[3] Identity
[4] Coverage
[5] *de novo* annotation

گرچه بسته‌های نرم‌افزاری مورد استفاده در متاژنومیکس ویروسی به‌صورت رایگان در دسترس می‌باشند، تجزیه و تحلیل داده‌ها اغلب به متخصصین بیوانفورماتیک ماهر نیاز دارد؛ نیازی که مانع کاربرد آسان NGS در شناسایی ویروس‌ها و ویروئیدها می‌شود. اخیرا واسط گرافیکی با کاربرد آسان به‌نام SearchSmallRNA، در دسترس قرار گرفته است تا ژنوم‌های ویروسی را با درجه اطمینان بالا با استفاده از اطلاعات به‌دست آمده از توالی‌یابی آر.ان.ای کوچک به‌کمک NGS، بازسازی کند. بسته‌های تجاری از قبیل CLC genomics Workbench and Geneious یا پلت فرم‌های باز مثل Galaxy نیز واسط‌هایی با کاربرد آسان را فراهم ساخته و پارامترسازی[6] به کمک این ابزارها ساده‌تر می‌شود. در نهایت این تلاش‌ها، شناسایی ویروس‌ها و ویروئیدها از طریق تکنولوژی‌های NGS را تسریع خواهد نمود (Zhang et al. 2014).

به‌طور کلی تجزیه و تحلیل مقادیر بسیار زیاد داده‌های توالی‌یابی، مهمترین چالش تکنیکی بیشتر آزمایشگاه‌های تشخیصی است. تولید حدود یک میلیون تا یک میلیارد توالی بسته به پلت فرم مورد استفاده حاصل از تنها یک واکنش توالی‌یابی، سبب ایجاد یک نیاز اساسی در جهت تغییر در مسیر ذخیره‌سازی، دسته‌بندی و دستورزی داده‌ها شده است (de Andrade and Vaslin 2014). تمام مطالعاتی که تاکنون در زمینه تشخیص ویروس‌ها و ویروئیدها صورت گرفته است، از یک سری مراحل واکاوی عمومی داده پیروی می‌کنند. ابتدا توالی‌های با کیفیت پایین حذف شده و در مرحله بعد پیرایش خوانش‌ها انجام می‌شود. سپس توالی‌ها با توالی‌های شناخته شده موجود در پایگاه‌های اطلاعاتی با استفاده از روش جستجوی تشابه توالی[7] با ابزارهایی مانند BLASTN و BLASTX مقایسه می‌شوند (Khan et al. 2009). درنهایت، موفقیت در این مسیر به وجود یک توالی مشابه قابل ردیابی در پایگاه‌های اطلاعاتی بستگی دارد. با این وجود، برخی جستجوهای تشابه توالی برای ردیابی ویروس‌های جدید و نوظهور زمانی که کل ترانسکریپتوم هیچ‌گونه تشابه توالی قابل شناسایی با توالی‌های از قبل ثبت شده ندارد، با شکست مواجه شده است. در تحقیق صورت گرفته بر روی توالی‌یابی ویروس

[6] Parameterization
[7] Sequence similarity



[ Downloaded from mg.genetics.ir on 2025-07-18 ]
[ DOR: 20.1001.1.20084439.1400.16.4.1.5 ]

رگه قهوه‌ای کاساوا (Cassava brown streak virus= CBSV) با استفاده از پلت فرم تعیین توالی حرارتی 454، تقریبا ۱۵٪ توالی‌های تولید شده، با استفاده از ابزار BLASTN یا BLASTX با توالی‌های موجود در بانک ژن تطابق (match) نشان ندادند. این توالی‌های مجهول ممکن است مربوط به بیمارگرهای ناشناخته باشند (Wu et al. 2012). بنابراین همان‌طور که در بخش قبل ذکر شد، زمانی‌که تشابه توالی قابل شناسایی وجود نداشته باشد، استفاده از روش‌های جدید مستقل از همولوژی برای تایید منشاء تاکسونومیکی این توالی‌های مجهول، ضروری است.

ردیابی توالی نوکلئیک اسید مشتق شده از بیمارگر الزاماً تایید کننده این نیست که بیمارگر مسئول بروز علائم است. در برخی مطالعات صورت گرفته شواهدی ارائه شده که بین بیمارگرهای شناخته شده و بیماری ارتباط وجود دارد. به‌عنوان مثال (2010) Bolduc et al. ژنوم کامل یک ویروس منفرد را در یک گیاه محک پیدا کردند، زمانی که از روی یک گونه گیاه زینتی دارای علائم ویروسی، مایه زنی شده بود. اگرچه این کار فرضیه کخ را کامل نکرده است، مدرکی فراهم می‌کند که طبق آن یک بیماری قابل انتقال با علائم مشابه ویروس وابسته به حضور ژنوم کامل یک کوکوموویروس[1] جدید می‌باشد. مطالعات دیگر تاکنون حضور چندین ویروس را به‌صورت آلودگی مخلوط نشان داده اند. در مطالعه (2011) Studholme et al. توالی‌هایی از سه ویروس جدید اضافی در گیاهان سیب‌زمینی شیرین آلوده به برخی ویروس‌ها که به‌عنوان عوامل بیماری‌زا در آن میزبان شناخته می‌شوند، کشف شد. تجزیه و تحلیل بیشتر، حضور ژنوم‌های ویروسی کامل را آشکار ساخت، در حالی که اهمیت این توالی‌های ویروسی جدید و نقش آن‌ها در بیماری چندان مشخص نیست.

تکنولوژی‌های توالی‌یابی با توان عملیاتی بالا احتمالاً بر اساس تنها اطلاعات توالی‌یابی، باعث تسریع در تشخیص بیمارگرها به‌ویژه در مورد ویروس و ویروئیدهای شناخته شده و توصیف شده می‌گردند. اما این امر به نوبه خود نوعی معمای بالقوه را برای نام‌گذاری و تاکسونومی ایجاد می‌کند، چون غیر از توالی، هیچ‌گونه اطلاعات دیگر (مانند ارتباط با بیماری، مورفولوژی پیکره، اندازه پیکره، ناقل، دامنه میزبانی و ... ) از این توالی‌ها به‌دست نمی‌آید. به علاوه در این موارد فرضیه کخ به ندرت کامل می‌شود و در برخی موارد تکمیل آن ممکن نیست. شاید در مورد ویروس‌ها و ویروئیدهایی که به‌طور فرضی شناسایی شده‌اند، افزودن یک پیشوند یا پسوند به نام‌هایشان، برای تاکید بر این نکته که ویروس یا ویروئید صرفا بر اساس توالی توصیف شده است، ضروری باشد. اطلاعات بیشتردر مورد این ویروس‌ها و ویروئیدها، با تحقیقات گسترده‌تر، افزایش خواهد یافت (de Andrade and Vaslin 2014).

**نتیجه‌گیری کلی**

تعیین کل محتوای نوکلئیک اسید در یک نمونه بیولوژیک با استفاده از تکنولوژی‌های NGS، ابزار قدرتمندی برای بیماری‌شناسان گیاهی در تشخیص و شناسایی ویروس‌ها و ویروئیدها فراهم می‌سازد. تکنولوژی‌های NGS احتمالاً سبب دگرگونی در خدمات بازرسی و قرنطینه خواهند شد، به گونه‌ای که یک سیستم راهنمای دقیق، سریع و کامل را برای یافتن ویروس‌ها و ویروئیدها در نمونه‌های گیاهی فراهم می‌سازند. برخلاف تکنیک‌های رایج از قبیل الایزا، PCR و ریزآرایه، رهیافت‌های متاژنومیکس نیاز به داشتن اطلاعات قبلی در مورد بیمارگرها ندارند. از سال ۲۰۰۹، شناسایی موفق ویروس‌ها و ویروئیدهای متنوع از طریق رهیافت‌های متاژنومیکس در گیاهان صورت گرفته است. اطلاعات به‌دست آمده نشان می‌دهند که در میان استراتژی‌های ایجاد شده به‌منظور غنی‌سازی خوانش‌های اختصاصی ویروس‌ها و یا ویروئیدها در NGS، توالی‌یابی با دقت و عمق زیاد مولکول‌های آر.ان.ای کوچک کل، موثرترین روش برای شناسایی ویروس‌های دی.ان.ای دار و آر.ان.ای دار و همچنین ویروئیدها بوده است (Studholme et al. 2011). در سال‌های اخیر، تجزیه و تحلیل نمونه‌های گیاهی به‌وسیله NGS و الگوریتم‌های محاسباتی در شناسایی دو ویروئید جدید و ٤٩ ویروس جدید از ۲۰ خانواده ویروسی شناخته شده، مؤثر بوده است. با وجود پراکندگی گسترده ویروئیدها در سیب‌زمینی، گونه‌های باغی و درختان میوه، تاکنون این بیمارگرها در تعداد زیادی از تک لپه‌ای‌ها، محصولاتی که میوه کوچک دارند و یا در حیوانات، یافت نشده‌اند.

[1] cucumovirus



[ Downloaded from mg.genetics.ir on 2025-07-18 ]

[ DOR: 20.1001.1.20084439.1400.16.4.1.5 ]

چالش‌های فراوانی در کاربرد تکنولوژی‌های NGS در مورد کشف بیمارگرها وجود دارد. اول، ایجاد و توسعه واسطهای نرم‌افزاری با کاربرد آسان مانند SearchSmallRNA که در دسترس عموم قرار گیرند و به اطلاعات انفورماتیک اندک نیاز داشته باشند، کاربرد تکنولوژی‌های NGS را در شناسایی و تشخیص ویروس‌ها و ویروئیدها تسهیل خواهد کرد. دوم، فرضیات کخ در مورد چندین ویروس و ویروئید کشف شده از طریق رهیافت‌های متاژنومیکس، هنوز کامل نشده‌است. اما به‌علت اینکه تولید مولکول‌های siRNA اختصاصی ویروس و ویروئید، نشان دهنده القای سیستم ایمنی میزبان در یک آلودگی فعال است، شناسایی ویروس‌ها و ویروئیدها از طریق توالی‌یابی مولکول‌های آر.ان.ای کوچک مدرکی برای تکثیر آن‌ها در میزبان فراهم می‌کند (Ding et al. 2012). برای بیماری شناسان گیاهی، این مطالعات ابزار مناسبی برای کشف ویروس‌ها و ویروئیدها هستند و شاید پیش‌بینی مشکلات آینده که می‌تواند گیاهان زیر کشت را تهدید کند. در واقع می‌توان با استفاده از داده‌های متاژنومیکس، دستیابی به یافته‌های مهمی از جمله شناسایی بیمارگرهای بالقوه برای امنیت بهتر غذایی و نیز افزایش اطلاعات در مورد برخی ویروس‌هایی که با آلوده‌سازی گیاهان اهلی می‌توانند به‌عنوان تهدید جدی محسوب شوند، را فراهم می‌آورد.

## منابع

Abia ALK, Alisoltani A, Keshri J, and Ubomba-Jaswa E (2018) Metagenomic analysis of the bacterial communities and their functional profiles in water and sediments of the Apies River, South Africa, as a function of land use. Science of the Total Environment 616:326-334.

Adams IP, Glover RH, Monger WA, Mumford R (2009) Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology. Molecular Plant Pathology 10:537-545.

Adams IP, Miano DW, Kinyua ZM, Wangai A. (2013) Use of next-generation sequencing for the identification and characterization of Maize chlorotic mottle virus and Sugarcane mosaic virus causing maize lethal necrosis in Kenya. Plant Pathology 62:741-49.

Alisoltani A, Ebrahimi S, Azarian S, Hematyar M (2015a) Parallel consideration of SSRs and differentially expressed genes under abiotic stress for targeted development of functional markers in almond and related prunus species. Science of Horticulture 198:462-472.

Alisoltani A, Fallahi H, Shiran B, Alisoltani A (2015b) RNA-Seq SSRs and small RNA-Seq SSRs: New approaches in cancer biomarker discovery. Gene 560:34-43.

Alisoltani A, Manhanzva MT, Potgieter M, et al. (2020) Microbial function and genital inflammation in young South African women at high risk of HIV infection. Microbiome 8:165.

Al Rwahnih M, Dave A, Anderson MM, Rowhani A, Uyemoto JK, Sudarshana MR (2013) Association of a DNA virus with grapevines affected by red blotch disease in California. Phytopathology 103:1069-76.

Altschul SF, Madden TL, Schäffer AA, Zhang J (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic acids research 25:3389-3402.

Aziz R, Bartels D, Best A, DeJongh M (2008) The RAST Server: rapid annotations using subsystems technology. BMC genomics 9:75.

Barzon L, Lavezzo E, Militello V, Toppo S (2011) Applications of next-generation sequencing technologies to diagnostic virology. International Journal of Molecular Science 12:7861–7884.

Boisvert S, Laviolette F, Corbeil, J. Ray (2010): Simultaneous assembly of reads from a mix of high-throughput sequencing technologies. Journal of Computational Biology 17:1519-1533

Bolduc F, Hoareau C, St-Pierre P, Perreault JP (2010) In-depth sequencing of the siRNAs associated with peach latent mosaic viroid infection. BMC Molecular Biology 11:16.

Bronzato Badial A, Sherman D, Stone A, Gopakumar A, Wilson V, Schneider W, King J (2018) Nanopore sequencing as a surveillance tool for plant pathogens in plant and insect tissues. Plant Disease 102:1648-1652.

Chaisson MJ, Pevzner PA (2008) Short read fragment assembly of bacterial genomes. Genome research 18:324-330. Cox-Foster DL, Conlan S, Holmes EC, Palacios G. (2007) A metagenomic survey of microbes in honey bee colony collapse disorder. Science 318:283-287.

Daneshpour MA, Fallah MS, Eshraghi P (2014) Revolution of DNA sequencing method from the past until Today. Modares Journal of Medical Sciences: Pathobiology 2014 16:1-13 (In Farsi).

de Andrade RRS, Vaslin MFS (2014) SearchSmallRNA: a graphical interface tool for the assemblage of viral genomes using small RNA libraries data. Virology Journal 11:45. Delwart EL (2007) Viral metagenomics. Reviews in Medical Virology 17:115-31.

Ding S-W, Lu R (2011) Virus-derived siRNAs and piRNAs in immunity and pathogenesis. Current Opinion in Virology 1:533-544.

Ding SW, Wang Y, Cao M, Ramachandran V (2012) Development of next-generation technologies for the diagnosis and identification of citrus viruses and viroids. Current Opinion in Virology 1: 533-544.

Eddy SR (1998) Profile hidden Markov models. Bioinformatics 14:755-763.



[ Downloaded from mg.genetics.ir on 2025-07-18 ]
[ DOR: 20.1001.1.20084439.1400.16.4.1.5 ]

Förstner KU (2009) Computational analysis of metagenomic data: delineation of compositional features and screens for desirable enzymes. PhD Thesis, 112p.
Grisoni M, Marais A, Filloux D, Saison A, Faure C, Julian C, Theil S, Contreras S, Teycheney PY, Roumagnac P (2017) Two novel Alphaflexiviridae members revealed by deep sequencing of the Vanilla (Orchidaceae) virome. Archives of Virology 162:3855-3861.
Gnerre S, MacCallum I, Przybylski D, Ribeiro FJ (2011) High-quality draft assemblies of mammalian genomes from massively parallel sequence data. Proceedings of the National Academy of Sciences 108:1513-1518.
Goll J, Rusch DB, Tanenbaum DM, Thiagarajan M, et al. (2010) METAREP: JCVI metagenomics reports—an open source tool for high-performance comparative metagenomics. Bioinformatics 26:2631-2632.
Gu YH, Tao X, Lai XJ, Wang HY (2014) Exploring the polyadenylated RNA virome of sweet potato through high-throughput sequencing. PLOS ONE 9:e98884.
Guindon S, Gascuel O (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic biology 52:696-704
Hach F, Hormozdiari F, Alkan C, Birol I (2010) mrsFAST: a cache-oblivious algorithm for short-read mapping. Nature methods 7:576-577
Handelsman J (2004) Metagenomics: application of genomics to uncultured microorganisms. Microbiology and Molecular Biology Reviews 68:669-685
Handelsman J, Rondon MR, Brady SF, Clardy J (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chemical biology 5:R245-R249
Homer N, Merriman B, Nelson SF (2009) BFAST: an alignment tool for large scale genome resequencing. PLoS One 4:e7767
Katsiani A, Maliogka VI, Katis N, Svanella-Dumas L, Olmos A, Ruiz-García AB, Marais A, Faure C, Theil S, Lotos L (2018) High-throughput sequencing reveals further diversity of Little cherry virus 1 with implications for diagnostics.Viruses 10:385.
Khan Z, Bloom JS, Kruglyak L, Singh M (2009) A practical algorithm for finding maximal exact matches in large sequence datasets using sparse suffix arrays. Bioinformatics 25:1609-1616.
Kircher M, Kelso J (2010) High-throughput DNA sequencing—concepts and limitations. Bioessays 32:524-536.
Klindworth A, Pruesse E, Schweer T, Peplies J (2012) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Research 41: 1-11.
Kreuze JF, Perez A, Untiveros M, Quispe D (2009) Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: a generic method for diagnosis, discovery and sequencing of viruses. Virology 388:1-7.
Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biology 10:R25.
Laserson J, Jojic V, Koller D (2011) Genovo: de novo assembly for metagenomes. J. Comput. Biol 18: 429-443
Levy SE and Myers RM (2016) Advancements in next-generation se- quencing. Annual Review of Genomics and Human Genetics 17:95-115.
Li H, Handsaker B, Wysoker A, Fennell T (2009a) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078-2079.
Li R, Yu C, Li Y, Lam TW (2009b) SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics 25:1966-1967.
Li Y, Lu JF, Han YH, Fan XX, et al. (2013) RNA interference functions as an antiviral immunity mechanism in mammals. Science 342:231-234.
Loconsole G, Saldarelli P, Doddapaneni H, Savino V, Martelli GP, Saponari M (2012). Identification of a single-stranded DNA virus associated with citrus chlorotic dwarf disease, a new member in the family Geminiviridae. Virology 432: 162-172.
Logan-Klumpler FJ, De Silva N, Boehme U, Rogers MB (2012) GeneDB—an annotation database for pathogens. Nucleic acids res. 40, D98-D108.
Lorenz P and Eck J (2005) Metagenomics and industrial applications. Nature Reviews Microbiology 3:510-516.
Lorenzi HA, Hoover J, Inman J, Safford T (2011) TheViral MetaGenome Annotation Pipeline (VMGAP): an automated tool for the functional annotation of viral Metagenomic shotgun sequencing data. Standards in genomic sciences 4:418-429.
Markowitz VM, Mavromatis K, Ivanova NN, Chen I (2009) IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics 25:2271-2278
Mehetre GT, Leo VV, Singh G, Sorokan A, Maksimov I, Yadav MK, Upadhyaya K, Hashem A, Alsaleh AN, Dawoud TM, Almaary KS, Singh BP (2021) Current Developments and Challenges in Plant Viral Diagnostics: A Systematic Review. Viruses 13:412.
Metzker ML (2010) Sequencing technologies-the next generation. Nature Reviews Genetics 11:31-46.
Mokili JL, Rohwer F, Dutilh BE (2012) Metagenomics and future perspectives in virus discovery. Current Opinion in Virology 2:63–77.
Monger WA, Alicai T, Ndunguru J, Kinyua ZM (2010) The complete genome sequence of the Tanzanian strain of Cassava brown streak virus and comparison with the Ugandan strain sequence. Archives of Virology 155:429–33.
Mumo NN, Mamati GE, Ateka EM, Rimberia FK, Asudi GO, Boykin LM, Machuka EM, Njuguna JN, Pelle R, Stomeo F (2020) Metagenomic analysis of plant viruses associated with Papaya Ringspot disease in Carica papaya L. in Kenya. Frontiers in Microbiology 11:205.
Mushtaq Z, Prasad KP, Qayoom U (2020) Nanopore Sequencing for diagnosis and resistance profiling of pathogens. Biotica Research 2:908-911.

[ Downloaded from mg.genetics.ir on 2025-07-18 ]
[ DOR: 20.1001.1.20084439.1400.16.4.1.5 ]