

نسل جدید روش‌های توالی‌یابی و کاربردهای آن

میثم مصلاهی^۱، حامد میرزایی^۲، میگانوش سیمونیان^۱، دکتر مجید خیرالهی^۳

مقاله مروری

چکیده

توالی‌یابی DNA یک روش آزمایشگاهی است که برای تعیین توالی یک مولکول DNA استفاده می‌شود و خود شامل هر روش یا تکنولوژی می‌باشد که برای تعیین ترتیب چهار باز آدنین، گوانین، سیتوزین و تیمین در یک رشته‌ی DNA به کار برده می‌شود. توالی‌یابی DNA ممکن است برای تعیین توالی ژن‌های ویژه، مناطق بزرگ ژنومی، کل کروموزوم و کل ژنوم به کار برده شود. امروزه، با پیشرفت‌های قابل ملاحظه‌ای که در بیولوژی مولکولی به وجود آمده است، روش‌های سنتی توالی‌یابی (به طور عمده تکنیک Sanger) که زمان‌بر و گران می‌باشند، دیگر پاسخگوی نیاز محققان نیستند. در نتیجه، افزایش تقاضا برای کاهش هزینه‌ی توالی‌یابی، به توسعه‌ی تکنولوژی‌های توالی‌یابی با توان عملکردی بالا (نسل جدید توالی‌یابی) منجر شد که قادر به توالی‌یابی هزاران یا میلیون‌ها توالی به صورت هم‌زمان می‌باشد. نسل جدید توالی‌یابی، قادر به توالی‌یابی سریع قطعات بزرگ DNA در طول ژنوم است و با حداقل ابزارها، قادر به تولید گیگا باز داده در هر مرحله‌ی توالی‌یابی می‌باشد. روش‌های نسل جدید توالی‌یابی، طیف وسیعی از کاربردها نظیر توالی‌یابی کل ژنوم، توالی‌یابی از نو، توالی‌یابی RNA برای کاربردهایی مثل بررسی ترانسکریپتوم و RNAهای کوچک، بررسی متیلاسیون و بررسی برهم‌کنش نوکلئیک اسید و پروتئین را شامل می‌شوند.

واژگان کلیدی: توالی‌یابی، نسل جدید توالی‌یابی، تکنولوژی ۴۵۴، تکنولوژی Solex، Sequencing by oligonucleotide ligation and detection

ارجاع: مصلاهی میثم، میرزایی حامد، سیمونیان میگانوش، خیرالهی مجید. نسل جدید روش‌های توالی‌یابی و کاربردهای آن. مجله دانشکده پزشکی اصفهان

۱۳۹۴؛ ۳۳ (۳۶۸): ۲۴۶۹-۲۴۸۰

مقدمه

Avery و همکاران در سال ۱۹۴۴ نشان دادند که دئوکسی ریبونوکلئیک اسید (DNA یا Deoxyribonucleic acid) را به عنوان ماده‌ی وراثتی مطرح نمود (۱). ساختار هلیکس دو رشته‌ای DNA که ترکیبی از چهار نوکلئوتید می‌باشد، در سال ۱۹۵۳ توسط Watson و Crick مشخص شد، که خود منجر به ارائه‌ی نظریه‌ی سنترال دوگما (Central dogma) در زیست‌شناسی مولکولی شد (۲). در اغلب موارد، DNA ژنومی به صورت گونه‌ای و فردی تعریف می‌شود و از توالی‌های آن برای بررسی عملکرد سلول و رمزگشایی راز و رمز موجود زنده استفاده می‌گردد. تکنیک‌های توالی‌یابی قادر هستند به زیست‌شناسان در طیف وسیعی از کاربردها مانند کلونینگ، پیدا کردن ژن‌های پاتوژن و مطالعات مقایسه‌ای کمک کنند. تکنیک‌های توالی‌یابی نسل جدید با توان عملیاتی بالا، یک دریچه‌ی جدید در

تحقیقات زیست‌پزشکی باز نمودند. همچنین، وجود نسل جدید دستگاه‌های توالی‌یابی، انقلاب عظیمی در ژنتیک ایجاد نموده‌اند و قادر به توالی‌یابی ژنوم انسان به قیمت ۱۰۰۰ دلار و حتی کمتر در طی چند روز با کیفیت بالا می‌باشند (۳).

تا چند سال پیش، برای توالی‌یابی از روش Sanger (روش دی‌دئوکسی) استفاده می‌شد. این روش، برای اولین بار در سال ۱۹۷۷ ارائه گردید. در همان سال، Maxam و Gilbert روش توالی‌یابی با تجزیه‌ی شیمیایی را مطرح نمودند (۴). روش Sanger، مجموعه‌ای از قطعات DNA را تولید می‌کند که اندازه‌ی آنها تنها در یک نوکلئوتید با یکدیگر تفاوت دارد (۵). در ابتدا، از ژل پلی‌اکریل آمید برای جداسازی این قطعات DNA استفاده می‌شد، اما به مرور زمان، ستون‌های کوچک جایگزین ژل‌های بزرگ شد. به کمک این ستون‌ها، می‌توان قطعات DNA را در ۲-۳ ساعت از یکدیگر جدا نمود. این

۱- دانشجوی کارشناسی ارشد، مرکز تحقیقات بیماری‌های ارثی کودکان، پژوهشکده‌ی پیش‌گیری اولیه از بیماری‌های غیر واگیر و گروه ژنتیک و بیولوژی مولکولی، دانشکده‌ی پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران

۲- دانشجوی دکتری، مرکز تحقیقات بیوتکنولوژی، دانشگاه علوم پزشکی مشهد، مشهد، ایران

۳- دانشیار، مرکز تحقیقات بیماری‌های ارثی کودکان، پژوهشکده‌ی پیش‌گیری اولیه از بیماری‌های غیر واگیر و گروه ژنتیک و بیولوژی مولکولی، دانشکده‌ی پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران

Email: mkheirollahi@med.mui.ac.ir

نویسنده‌ی مسؤول: دکتر مجید خیرالهی

و تجاری‌سازی تکنولوژی NGS می‌باشد. این شرکت، در اواخر دهه‌ی ۱۹۹۰ برای اولین بار از تکنولوژی موسوم به پایروسکوئینسینگ (Pyrosequencing) برای تعیین توالی DNA در مقیاس وسیع استفاده کرد. در این سیستم، نمونه‌ی DNA به قطعاتی با طول حدود ۱۰۰ جفت باز شکسته می‌شود. سپس، دو آداپتور کوتاه به انتهای قطعات متصل می‌شود. این آداپتورها، محل اتصال پرایمرهای لازم برای تکثیر و توالی‌یابی هستند. این قطعات، با ذرات کوچکی که پوششی از جنس استرپتایدین در سطح خود دارند، مخلوط می‌شوند و به کمک آداپتور B خود که در انتهای □ ۵ بیوتینه است، به این ذرات متصل می‌گردند. آن‌گاه، قطعات دو رشته‌ای متصل به ذرات، دناتور و تک رشته‌ای می‌گردند. این مخلوط، به اندازه‌ی کافی رقیق است؛ به طوری که به هر ذره‌ی منفرد، تنها یک رشته‌ی DNA متصل می‌باشد. سپس، با افزودن ترکیبات لازم برای تکثیر قطعات DNA متصل به ذرات، این مخلوط به امولسیون تبدیل می‌شود. به این ترتیب، هر ذره با رشته‌ی DNA متصل به آن، در یک قطره‌ی امولسیون به دام می‌افتد و واکنش Polymerase chain reaction (PCR) به طور مستقل در هر قطره انجام می‌گردد (Emulsion PCR).

در مرحله‌ی بعد، به منظور تعیین توالی، ذرات حاوی DNA بر روی صفحه‌ی سلیکونی متشکل از ۴۰۰۰۰۰۰ چاهک منظم با حجمی در حد پیکولیترا پخش می‌شوند (۱۱). اندازه‌ی کوچک چاهک، باعث می‌شود تا در هر چاهک بیش از یک ذره قرار نگیرد (شکل ۱). سپس آنزیم‌های لازم یعنی DNA پلیمراز، سولفوریلاز و لوسیفراز به چاهک افزوده می‌شوند. چاهک‌ها به طور مجزا پی در پی در معرض Deoxycytidine triphosphate (dCTP)، (Deoxythymidine triphosphate) dTTP، و Deoxyguanosine triphosphate (dATP) قرار می‌گیرند و بین هر بار افزودن سوبسترای دنوکسی نوکلئوتید، یک مرحله‌ی شستشو انجام می‌شود. بدیهی است که ورود هر نوکلئوتید به رشته‌ی در حال گسترش، وابسته به حضور باز مکمل در الگو می‌باشد. با ورود هر نوکلئوتید مکمل، یک مولکول پیروفسفات رها می‌شود. آنزیم سولفوریلاز، این مولکول پیروفسفات را با استفاده از آدنوزین فسفوسولفات (Adenosine phosphosulfate یا APS) به ATP تبدیل می‌کند. آنزیم لوسیفراز نیز به کمک ATP حاصل شده، سوبسترای خود را که لوسیفیرین می‌باشد، به اکسی لوسیفیرین تبدیل می‌کند. این واکنش همراه با ساطع شدن نور می‌باشد. نور ساطع شده توسط دوربین‌های متصل به صفحه‌ی سلیکونی ثبت و به صورت پیک در یک گراف نمایان می‌شود (شکل ۱).

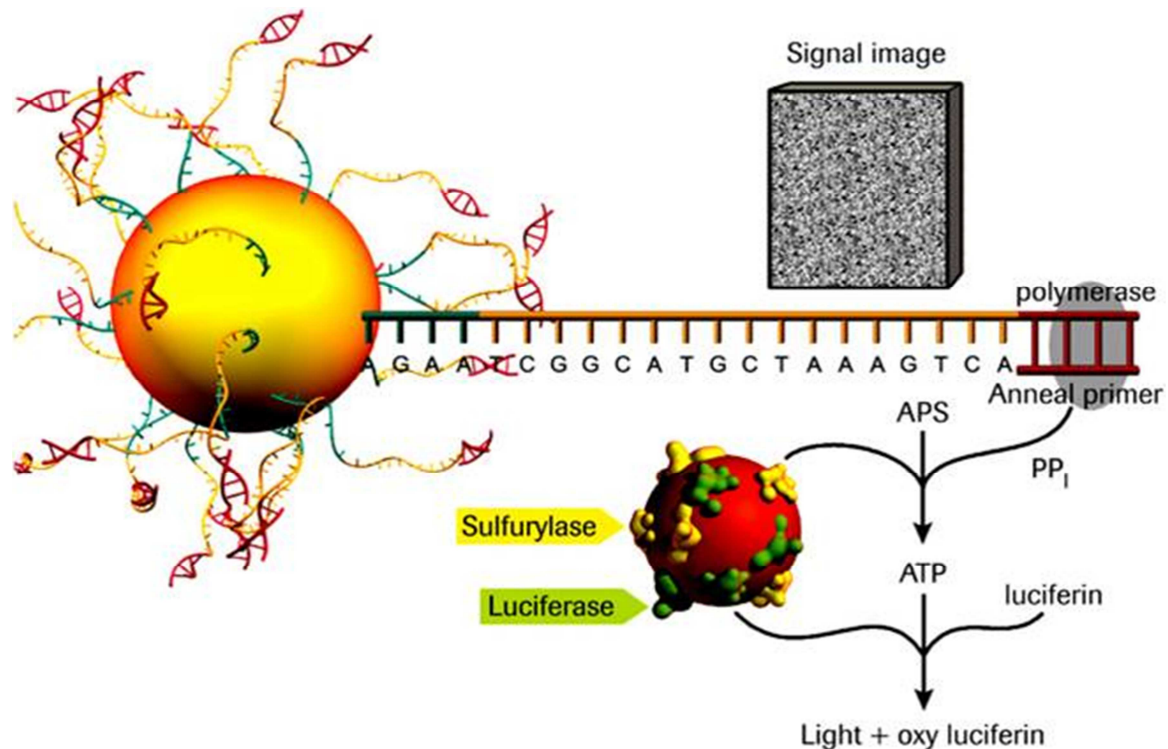
ستون‌ها، قابلیت استفاده‌ی مجدد را دارند و جدا کردن قطعات بین ۷۰۰-۸۰۰ جفت باز را امکان‌پذیر می‌سازند؛ یعنی ظرفیتی مشابه با ژل پلی‌اکریل آمید دارند، اما کار با آن‌ها به مراتب ساده‌تر می‌باشد (۶). با استفاده از نوکلئوتیدهای فلورسنت خاتمه دهنده‌ی زنجیره، پیشرفت مهم دیگری در توالی‌یابی DNA حاصل شد. به طور کلی، در این روش آخرین نوکلئوتید هر یک از قطعات DNA با یک رنگ فلورسنت نشانه‌گذاری می‌شود. بنا بر این، هر قطعه‌ی DNA دارای اندازه و رنگ منحصر به فردی خواهد بود (۷). هم‌زمان با جدا شدن این قطعات بر روی ستون که بر اساس اندازه‌ی آن‌ها انجام می‌شود، حس‌گرهای فلورسنت، رنگ هر قطعه‌ی DNA را شناسایی می‌کنند. به این ترتیب، یک ستون، یک توالی ۶۰۰-۸۰۰ جفت بازی را در کمتر از ۳ ساعت بر اساس اندازه‌ی جداسازی و رمزگشایی می‌کند. دستگاه‌های توالی‌یابی خودکار که از نوکلئوتیدهای فلورسنت استفاده می‌کنند، به گونه‌ای توسعه یافته‌اند که دارای ۳۸۴ ستون مجزا کننده هستند و قادرند در یک اجرای سه ساعته، ۲۰۰ کیلو جفت باز DNA را توالی‌یابی کنند (۸). مجموعه‌ای از صد دستگاه می‌تواند اطلاعاتی معادل توالی ژنوم انسان را تنها در دو ماه ایجاد نمایند. توالی‌های تولید شده توسط نسل جدید دستگاه‌های تعیین توالی، در اصطلاح Read (خوانش) نامیده می‌شوند.

به طور کلی، توالی‌یابی قطعات DNA با استفاده از تکنیک‌های جدید توالی‌یابی (Next generation sequencing یا NGS)، بر خلاف توالی‌یابی به روش Sanger، نیازی به مراحل همسانه‌سازی ندارد؛ این مسأله، زمان و هزینه‌ی توالی‌یابی را به شدت کاهش داده است. از دیگر مزیت‌های این نوع توالی‌یابی، می‌توان به تولید میلیون‌ها توالی به طول متوسط ۷۰۰-۳۰۰ جفت باز (بسته به هدف مورد نظر و نوع سیستم استفاده شده) تنها در طی چند روز با دقت و صحت بالا (توالی‌یابی در مقیاس وسیع) اشاره نمود. این ویژگی‌ها باعث شد تا این تکنولوژی جایگزین روش توالی‌یابی Sanger شود و اغلب آزمایشگاه‌ها بتوانند به طور مستقل از مراکز توالی‌یابی ژنومی بزرگ از آن بهره‌برداری کنند (۹).

در کنار مزایای ارزشمند تکنیک‌های NGS، آنالیز داده‌های حاصل از توالی‌یابی که طول کوتاهی دارند، بزرگ‌ترین چالش در این زمینه می‌باشد که نیازمند کامپیوترهای پیشرفته، نرم‌افزارهای خاص برای آنالیز داده‌ها، همچنین دانش و تخصص کافی است (۱۰). نسل جدید دستگاه‌های توالی‌یابی در اواخر دهه‌ی ۱۹۹۰ توسط شرکت‌های مختلف نظیر Illumina/Solex، Life science/Roch/454 و ABI/SOLiD ایجاد و طی سال‌های ۲۰۰۵-۰۶ تجاری شدند.

تکنولوژی Roch/454 Life science

شرکت Roch/454 Life science اولین مؤسسه‌ی پیشگام در تولید



شکل ۱. تعیین توالی با تکنولوژی ۴۵۴ (۵۷)

هر اجرای این دستگاه اندک است، استفاده از آن برای موجوداتی با اطلاعات ژنتیک محدود، مستلزم اجرای متعدد می‌باشد که از نظر اقتصادی چندان مقرون به صرفه نیست. در مقابل، این تکنیک ابزاری قدرتمند برای توالی‌یابی مجدد ژنوم‌های شناخته شده می‌باشد (۱۱). به عنوان مثال، این روش برای توالی‌یابی بخش‌هایی از ژن‌های یک شخص، به منظور تشخیص جهش‌های بیماری‌زا به خوبی عمل می‌کند. به منظور افزایش سرعت در این مورد، نوکلئوتیدها می‌توانند به همان ترتیبی که در توالی طبیعی شناخته شده وجود دارند، اضافه شوند. به این ترتیب، جهش از طریق عدم توانایی نوکلئوتید طبیعی برای اتصال در یک موقعیت خاص به سادگی تشخیص داده می‌شود (۱۰).

تکنولوژی Illumina/Solex

شرکت Solex که در سال ۲۰۰۶ دستگاه Genome analyzer را به بازار ارایه کرد که در سال ۲۰۰۷ توسط شرکت Illumina خریداری شد و از آن زمان به بعد با نام Illumina/Solex شناخته می‌شود. طول متوسط توالی‌های خوانده شده توسط این تکنولوژی، ۱۵۰-۳۷ جفت باز می‌باشد (۸). برای انجام توالی‌یابی به کمک این تکنولوژی، ابتدا DNA مورد نظر به قطعاتی با طول کمتر از ۸۰۰ جفت باز شکسته می‌شود. این قطعات، دارای انتهای صاف هستند و از دو طرف به آداپتورهایی متصل می‌شوند. فرآورده‌های حاصل، پس از

رونق‌ی و همکاری‌ها، در سال ۱۹۹۸ از پایه‌گذاران پاپروسکوئینسینگ بودند (۱۲). در آخرین نسخه‌ی ارایه شده از طرف شرکت Roch، این تکنولوژی می‌تواند تا ۱۴ گیگا باز (Giga bases; Gb) توالی با طول متوسط ۷۰۰ جفت باز تولید کند. طول بلند توالی‌های تولید شده توسط این روش و سرعت بالا، از مهم‌ترین ویژگی‌های این تکنولوژی می‌باشد. هر اجرای پاپروسکوئینسینگ بسیار سریع است؛ در واقع، این تکنیک می‌تواند تعداد زیادی اجرا را به طور هم‌زمان انجام دهد. به عنوان مثال، ۹۶ اجرای مختلف می‌تواند به طور هم‌زمان در یک پلیت ۹۶ چاهکی انجام شود، از آن جایی که کل فرایند اتوماتیک است، نیاز چندانی به توجه فرد آزمایش‌کننده ندارد. بدیهی است هر تکنولوژی در کنار مزایای خود، یک سری معایب نیز دارد و یکی از معایب این روش، درصد خطای به نسبت بالای آن در نواحی هموپلیمر DNA است. هزینه‌ی بالای توالی‌یابی با این روش در مقایسه با سایر روش‌های توالی‌یابی در مقیاس وسیع، عیب دیگر این تکنولوژی به شمار می‌رود که استفاده از آن را تا حدی محدود می‌کند (۱۳).

تکنولوژی ۴۵۴، توالی‌های بلندتری را در مقایسه با روش‌های دیگر توالی‌یابی در مقیاس وسیع تولید می‌کند. این مسأله، دقت و سر هم نمودن این توالی‌ها را به توالی‌های بلندتر را به طور قابل توجهی افزایش می‌دهد. با این وجود، چون تعداد خوانش‌های تولید شده در

۲۰۱۰، Illumina سیستم HiSeq را به بازار ارایه کرد که از راهبرد پیش‌گفته، برای تعیین توالی استفاده می‌کند، اما خروجی آن در مقایسه با نسخه‌های قبلی بسیار بیشتر است. این سیستم در ابتدا در هر اجرا قادر به خواندن ۲۰۰ Gb توالی بود که به ۶۰۰ Gb افزایش یافته است (۱۵). اگر چه طول کوتاه توالی‌های تولید شده توسط Illumina انجام آنالیزهای دقیق به ویژه در موجوداتی با اطلاعات ژنتیک بسیار کم را محدود می‌کند، اما قابلیت انجام توالی‌یابی در هر دو جهت مولکول DNA که به Paired-end sequencing معروف است، تا حد زیادی بر این مشکل غلبه کرده است؛ به طوری که امروزه اغلب با استفاده از این سیستم و در مواردی ترکیب دو سیستم ۴۵۴ و Illumina ژنوم بسیاری از موجودات در حال توالی‌یابی است (۸).

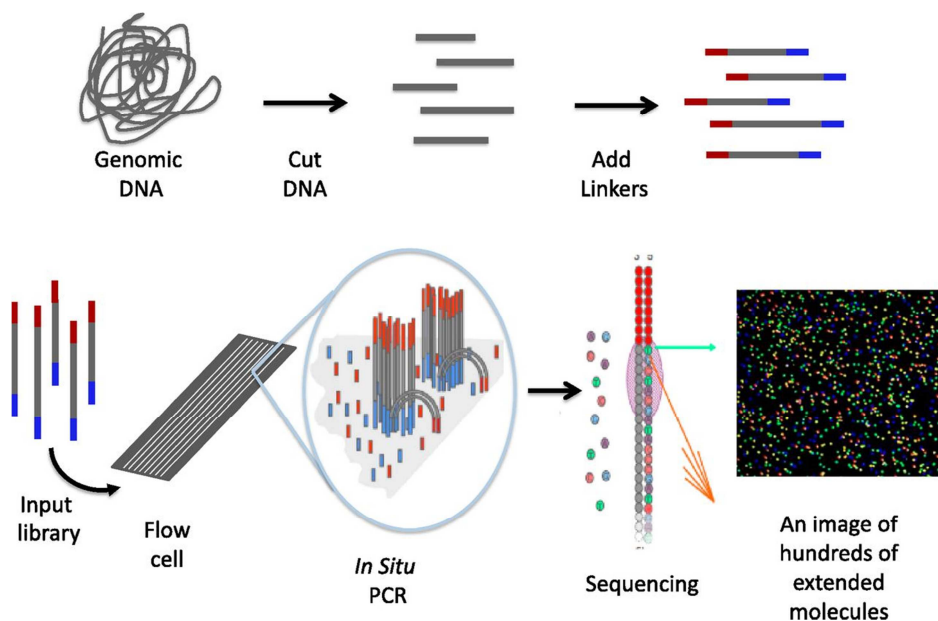
تکنولوژی ABI/SOLiD

شرکت Sequencing by oligonucleotide Ligation and detection (SOLiD) در سال ۲۰۰۶، تکنولوژی توالی‌یابی از طریق اتصال (Sequencing by ligation) را به بازار معرفی کرد و در همین سال نیز توسط شرکت Applied biosystems (ABI) خریداری شد. با استفاده از این تکنولوژی، توالی‌هایی به طول ۳۵-۷۵ جفت باز خوانده می‌شود. مراحل آماده‌سازی نمونه‌ی DNA، به طور تقریبی مشابه با سیستم توالی‌یابی ۴۵۴ است؛ به طوری که نمونه‌ی DNA مورد نظر به قطعات کوچک‌تری (حدود ۲۰۰ جفت باز) شکسته می‌شود و پس از اتصال آداپتور به انتهای قطعات حاصل، این قطعات از طریق Emulsion PCR تکثیر می‌گردند (۸).

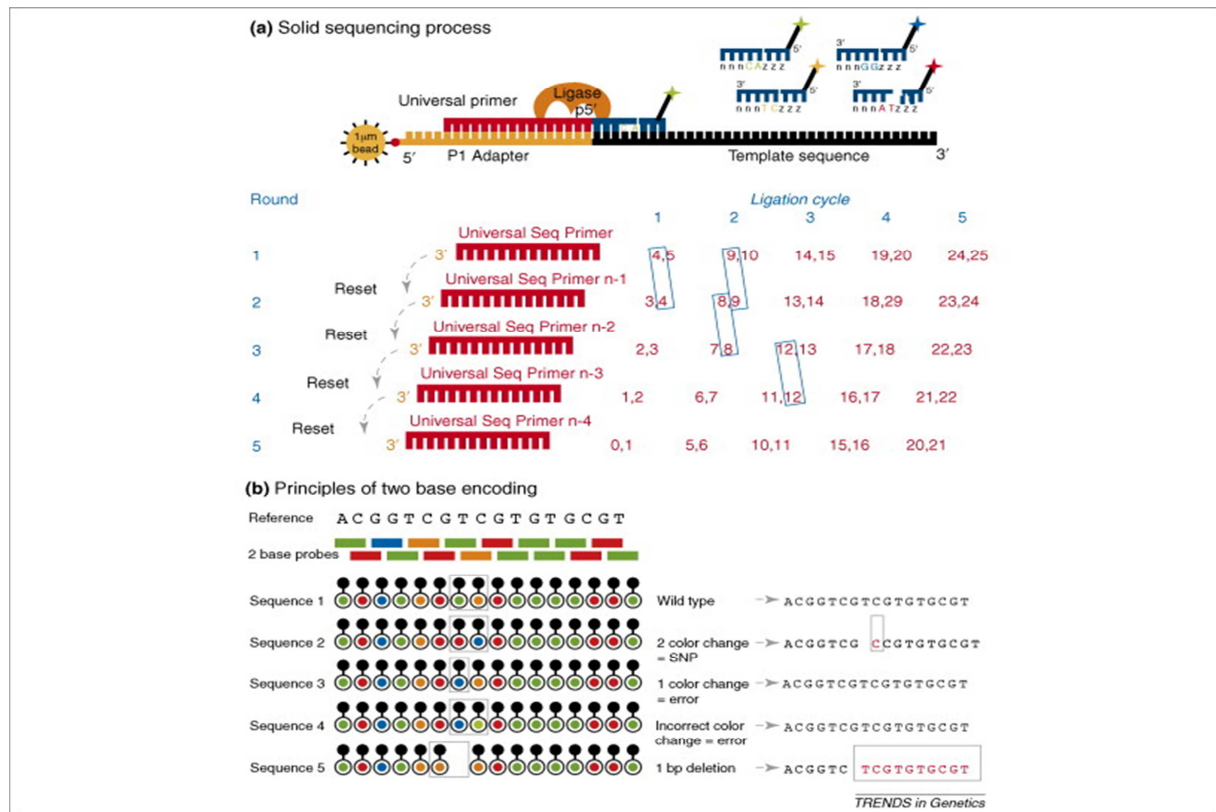
دنا توره شدن از یک انتها با پرایمرهایی متصل (و مکمل با توالی آداپتور) به یک سطح جامد به نام Flow cell هیبرید می‌شوند و بر روی سطح تثبیت می‌شوند و از انتهای دیگر، با توالی مکمل دیگر در سطح Flow cell هیبرید می‌شوند و به این ترتیب، یک ساختار پل مانند شکل می‌گیرد که الگوی تکثیر در واکنش PCR است (Bridge PCR). پس از انجام واکنش PCR، در حدود ۲۰۰۰ کپی از قطعات DNA دو رشته‌ای بر روی سطح جامد تولید می‌شوند که به آن‌ها در اصطلاح خوشه (Cluster) گفته می‌شود (شکل ۲).

در مرحله‌ی بعد، به منظور تعیین توالی قطعات، مولکول‌های DNA دناتور و خطی می‌شوند، پس از هیبریداسیون پرایمر مخصوص توالی‌یابی به توالی مکمل آداپتوری در انتهای هر رشته، Flow cell برای تعیین توالی در دستگاه مربوط آماده است. در هر چرخه‌ی توالی‌یابی، مخلوطی از چهار نوکلئوتید که هر کدام با رنگ فلورسنت مختلف نشانه‌گذاری شده‌اند (رنگ فلورسنت واجد یک گروه خاتمه دهنده‌ی سنتز زنجیره نیز می‌باشد)، به سطح Flow cell اضافه می‌شود. پس از شناسایی نوکلئوتید اضافه شده به رشته‌ی DNA در حال سنتز، رنگ فلورسنت به همراه گروه خاتمه دهنده از انتهای باز برداشته می‌شود و با اضافه شدن مخلوط نوکلئوتیدی به سطح، چرخه‌ی بعدی آغاز می‌گردد (شکل ۲).

به این ترتیب، در هر چرخه، تنها یک نوکلئوتید به رشته‌ی DNA در حال سنتز افزوده می‌شود (۱۴). بر خلاف تکنولوژی ۴۵۴، فرایند اضافه شدن بازها در این سیستم، به خوبی قادر به تعیین توالی DNA با دقت بالا در نواحی هموپلیمر و توالی‌های تکراری است. در اوایل سال



شکل ۲. تعیین توالی در تکنولوژی Illumina (۵۸)



شکل ۳. توالی‌یابی (ABI/SOLiD sequencing) Applied biosystems/Sequencing by oligonucleotide ligation and detection (۵۷)

Two base encoding نامیده می‌شود (شکل ۳).

صحت توالی‌های خوانده شده با استفاده از این تکنولوژی، در مقایسه با سیستم‌های دیگر در سطح بالایی است. به کمک این روش تعیین توالی، می‌توان بین خطای توالی‌یابی (Sequencing error) و چند شکلی تک نوکلئوتیدی (Single nucleotide polymorphism) تمایز قائل شد. این تمایز با در نظر گرفتن این مسأله است که خطای توالی‌یابی تنها در یک واکنش تعیین توالی قابل ردیابی است، اما چند شکلی در هر دو واکنش شناسایی می‌شود. از این رو، دقت توالی‌یابی در این سیستم در مقایسه با سیستم‌های دیگر در سطح بالایی است. در ابتدا، در هر اجرای دستگاه تا ۳ Gb توالی خوانده می‌شد، اما در دسامبر ۲۰۰۹، نسخه‌ی جدیدی از این تکنولوژی تحت عنوان Plus Plus 3TMSOLiD معرفی شد که قادر است در هر اجرا، بیش از ۶۰ Gb توالی را بخواند (۱۴).

تعیین توالی مولکول منفرد (Single molecule sequencing)

در حال حاضر، جدیدترین سیستم توالی‌یابی که به نسل سوم توالی‌یابی (Next- next- generation sequencing) معروف است، اولین بار توسط شرکت Helicos Bioscience معرفی شد و قادر به توالی‌یابی یک مولکول منفرد می‌باشد (۱۶). ویژگی‌های

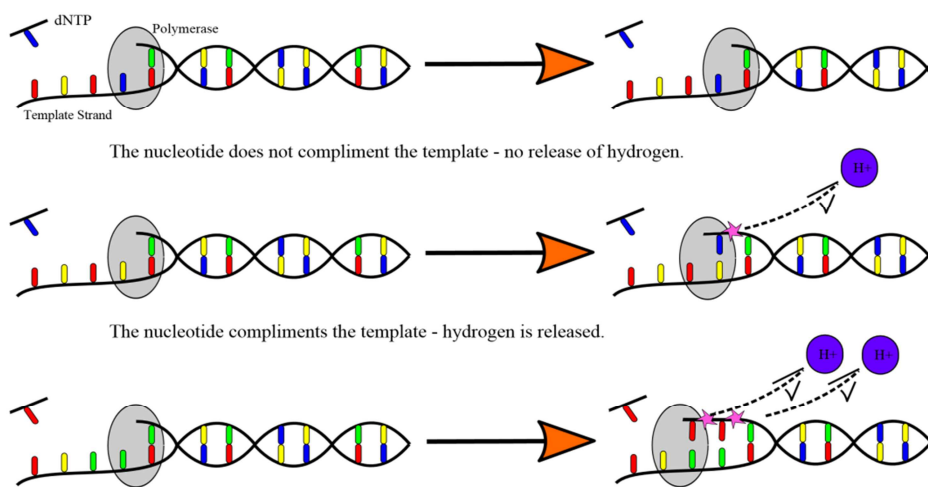
در مرحله‌ی تعیین توالی، پرایمر توالی‌یابی که مکمل توالی آداپتور در انتهای قطعات DNA است، به این قطعات متصل می‌شود، سپس یک سری اکتامرهای اولیگو نوکلئوتیدی نشان‌دار شده با رنگ‌های فلورسنت، برای اتصال به پرایمر با یکدیگر رقابت می‌کنند. در این اکتامر، دو باز وسط (بازهای ۴ و ۵) با یکی از چهار رنگ فلورسنت نشان‌دار شده است. به این ترتیب، پس از اتصال اکتامر به پرایمر این بازها شناسایی خواهند شد (۱۴). اکتامر اولیگو نوکلئوتیدی پس از باز پنجم شکسته می‌شود و رنگ فلورسنت نیز آزاد می‌گردد. سپس، با اتصال اکتامر دیگر چرخه‌های هیبریداسیون و اتصال تکرار می‌شوند تا بازهای ۹ و ۱۰ تعیین توالی شوند. در چرخه‌ی بعدی، بازهایی که در موقعیت‌های ۱۴ و ۱۵ قرار گرفته‌اند، توالی‌یابی می‌گردند. پس از یک سری چرخه‌های اتصال (Ligation)، قطعه‌ی مورد نظر با پرایمر مخصوص دیگری که مکمل موقعیت n-1 است، هیبرید می‌شود و دوباره، چرخه‌های اتصال اکتامر اولیگو نوکلئوتیدی به پرایمر تکرار می‌گردد تا بازهای دیگر توالی‌یابی گردند. پس از ۵ دور هیبریداسیون پرایمرهای توالی‌یابی مختلف با قطعه‌ی مورد نظر و تکمیل تکرار چرخه‌های اتصال برای هر اکتامر اولیگو نوکلئوتیدی نشان‌دار، هر باز با استفاده از دو پرایمر توالی‌یابی و در طی دو واکنش اتصال مستقل توالی‌یابی می‌شود. این روش تعیین توالی در اصطلاح

تحقیقات میکروبی و شناسایی پاتوژن‌ها، بسیار مناسب و کاربردی می‌باشد (۱۷).

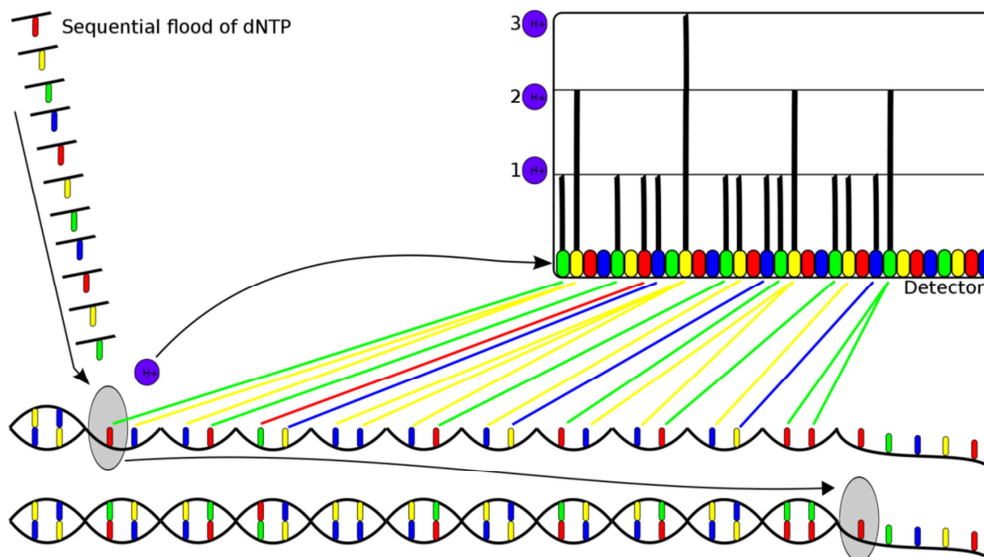
شرکت Ion Torrent در اواخر سال ۲۰۱۰، دستگاهی موسوم به Ion personal genome machine (Ion PGM) را روانه‌ی بازار کرد که نوع دیگری از دستگاه نسل سوم توالی‌یابی است. در این تکنولوژی، به ازای اضافه شدن هر نوکلئوتید توسط پلیمرز به مولکول DNA، یک پروتون آزاد می‌گردد و pH محیط واکنش تغییر می‌یابد. به این ترتیب، تغییر pH توسط حس‌گرهایی ردیابی می‌شود و اضافه شدن نوکلئوتید و نوع آن را ردیابی می‌کند (شکل ۴) (۱۶).

ممتاز این تکنولوژی در مقایسه با تکنیک‌های نسل دوم توالی‌یابی، عدم نیاز به تکثیر DNA با واکنش PCR می‌باشد که باعث افزایش دقت و کاهش زمان لازم برای توالی‌یابی شده است؛ همچنین، قابلیت شناسایی نوکلئوتید اضافه شده به زنجیره‌ی در حال ساخت، در زمان واقعی (Real time) و بلافاصله پس از اضافه شدن از امتیازات این تکنیک می‌باشد. طول متوسط توالی‌های خوانده شده توسط این سیستم، ۱۳۰۰ جفت باز است که از طول توالی‌های خوانده شده توسط تمام دستگاه‌های نسل دوم بیشتر است. با این وجود، این تکنولوژی برای آزمایشگاه‌های کلینیک و به ویژه

A



B



شکل ۴. A) اضافه شدن نوکلئوتیدها به رشته‌ی در حال سنتز و آزاد شدن پروتون‌ها. B) تعیین توالی به وسیله‌ی تکنولوژی شرکت Ion Torren، ردیابی هر نوکلئوتید به وسیله‌ی حس‌گرهای ویژه (۸)

کاربردهای NGS

سرعت بالا و هزینه‌ی پایین توالی‌یابی با استفاده از این تکنیک، در مقایسه با روش تعیین توالی Sanger، کاربردهای فراوانی برای آن ایجاد کرده است که به برخی از آن‌ها اشاره می‌شود (جدول ۱).

توالی‌یابی کل ژنوم به منظور دست یافتن به واریانت‌های مختلف مرتبط با بیماری‌ها

دست یافتن به واریانت‌های ژنتیک که افراد را نسبت به دیگر افراد جامعه مستعد ابتلا به یک بیماری می‌نماید، از ابتدا مورد توجه بوده است. مطالعه بر روی کل ژنوم از طریق توالی‌یابی، در ابتدا بر روی واریانت‌هایی که در جامعه شایع هستند ($MAF > 5\%$ یا Minor allele frequency) انجام می‌گرفت (۱۸)، اما با دست‌یابی به روش‌های NGS و توالی‌یابی کل ژنوم از طریق آن، امکان بررسی واریانت‌هایی با فراوانی پایین‌تر (۱ درصد $MAF >$) و بر روی تعداد افراد بیشتری فراهم شد، به عنوان مثال در طرح ۱۰۰۰ ژنوم، واریانت‌هایی با فراوانی کمتر از ۱ درصد، مورد بررسی قرار گرفتند (۲۰-۱۹).

توالی‌یابی مناطق هدف (Targeted sequencing)

اگر چه توالی‌یابی کل ژنوم با استفاده از NGS روش بهینه و جامع‌تری برای آنالیز ژنوم می‌باشد، اما این تکنیک در بسیاری از

آزمایشگاه‌ها تحقیقی و بالینی در دسترس نمی‌باشد (۲۱)، در نتیجه توالی‌یابی یک سری توالی‌های خاص که جهش یا تغییرات دیگر در آن‌ها موجب بیماری می‌شود، مورد توجه قرار گرفته است. امروزه، از پلت فرم‌های مختلف NGS به منظور Exome sequencing در بیماری‌هایی که چندین ژن مختلف مستعد برای جهش دارند و یا بدون نقاط Hot spot برای جهش می‌باشند، استفاده می‌شود (۲۳-۲۲، ۱۸). از این موارد، می‌توان به بیماری‌های نقص ایمنی و سرطان‌ها مانند کلورکتال و یا سینه اشاره نمود که می‌تواند حاصل جهش در یک سری ژن‌های هدف باشد و یا به عنوان مثال برای بررسی تغییر طول میکروساتلایت‌ها (Microsatellite instability by next-generation sequencing) یا msSINGS در سرطان کلورکتال اشاره نمود که هم در تشخیص نوع سرطان و هم در درمان آن بسیار مفید می‌باشد (۲۵-۲۴).

تعیین توالی از نو و اسمبلی (De novo sequencing and Assembly)

توالی‌یابی از نو، به مفهوم تولید توالی اولیه‌ی ژنومی یا ترانسکریپتومی برای موجوداتی است که اطلاعات ژنتیک محدودی از آن‌ها در دسترس می‌باشد (۲۶).

جدول ۱. کاربردهایی از روش‌های نسل دوم توالی‌یابی

سال	منبع	مثال کاربرد	کاربرد
۲۰۰۷	Velasco و همکاران (۵۹)	تعیین توالی از نو	ژنوم
۲۰۰۹	Huang و همکاران (۶۰)		
۲۰۱۰	Li و همکاران (۶۱)		
۲۰۰۶	Bentley (۲۶)	توالی‌یابی مجدد	
۲۰۰۸	Ossowski و همکاران (۲۷)		
۲۰۰۹	Denver و همکاران (۶۲)	توالی‌یابی کل ژنوم	
۲۰۰۹	Wang و همکاران (۲۸)	RNA-Seq	ترانسکریپتوم
۲۰۰۴	Impey و همکاران (۵۳)	پروفایل ژنومی از تغییرات هستون	اپی‌ژنتیک
۲۰۰۷	Mikkelsen و همکاران (۵۲)		
۲۰۰۸	Cokus و همکاران (۶۳)	متیلاسیون DNA	
۲۰۰۹	Costello و همکاران (۶۴)		
۲۰۰۶	Fierer و همکاران (۶۵)	پروفایل ژنومی از موقعیت نوکلئوزوم	
۲۰۰۶	Johnson و همکاران (۶۶)		
۲۰۰۷	Turnbaugh و همکاران (۶۷)		
۲۰۱۰	Qin و همکاران (۶۸)		
۲۰۰۶	Edward و همکاران (۵۵)	فلور میکروبی محیطی	متاژنوم
۲۰۰۷	Huber و همکاران (۵۴)		

دارند، از این رونوشت‌ها می‌توان به عنوان نشانگرهای تشخیصی و کمک کننده به تعیین نوع درمان نیز استفاده نمود که با توجه به کارآمدی NGS در تعیین توالی RNA کل، می‌توان از آن به عنوان یک ابزار مناسب در تشخیص و کمک کننده به درمان یاد نمود (۳۴، ۱۸).

بررسی تغییرات اپی ژنتیک

ارتباط بین پروتئین و DNA یک برهم‌کنش بیولوژیک است که نقش عمده‌ای در قابل دسترس بودن DNA و تنظیم بیان ژن‌ها دارد. در سال‌های اخیر، برای بررسی پروفایل پروتئین‌های متصل شده به DNA در مقیاس ژنومی و همچنین، تغییرات پروتئین‌های هیستونی و نوکلئوزوم‌ها، از تکنیک *Combining chromatin immunoprecipitation* (ChIP-seq) استفاده می‌شود (۳۵). به این ترتیب که پس از رسوب‌دهی ایمونوکروماتینی و تفکیک پروتئین‌ها، DNA به کمک تکنیک NGS توالی‌یابی می‌گردد. از این جمله، می‌توان به مطالعات انجام شده با استفاده از تکنولوژی ۴۵۴ و تعیین جایگاه اتصال برخی عوامل نسخه‌برداری به DNA با استفاده از سیستم توالی‌یابی Illumina اشاره کرد (۳۶). در طرح ENCODE با استفاده از ChIP-seq، ۱۲ نوع تغییر هیستونی در ۴۶ رده سلولی مختلف تعیین شد (۳۷). به منظور نقشه‌یابی الگوی متیلاسیون ژنوم، از تکنیک‌های مختلف NGS مانند *Methylated DNA immunoprecipitation* (MeDIP-Seq)، *MethylC-seq* و *Reduced representation bisulfite sequencing* (RRBS) (۳۸-۴۰) استفاده می‌شود که یکی از کاربردهای بالینی آن، در تعیین الگوی متیلاسیون در ژن‌های مستعدی مانند *MutL homolog 1* (MLH1) و *MutS Homolog 2* (MSH2) در سرطان کلورکتال می‌باشد (۴۱).

توالی‌یابی تک سلول (*Single-cell sequencing*)

بسیاری از روش‌های توالی‌یابی، نیاز به DNA یا RNA حاصل از 10^5 سلول دارند (۴۳-۴۲) که این امر، در مورد بافت‌های سخت توموری که بسیار هتروژن و همچنین همراه سلول‌های سالم می‌باشند، مسأله‌ساز می‌باشد. در نتیجه، تشخیص سلول‌های اولیه و در واقع، تشخیص کلون‌های سرطانی مشکل می‌شود (۴۴)؛ اما امروزه، از آن جا که NGS به DNA اولیه‌ی پایینی نیاز دارد، می‌توان به صورت تک سلول به توالی‌یابی کل ژنوم پرداخت و به تعداد کلون‌های ایجاد کننده‌ی سرطان پی برد (۴۵). به عنوان مثال، اولین بار Navin و همکاران، با جدا نمودن ۱۰۰ سلول از بافت سرطان سینه و انجام توالی‌یابی تک سلولی، نشان دادند که تومور از ۳ کلون اولیه تشکیل شده است که این امر، در زمینه‌ی تشخیص و انتخاب روش درمانی مفید می‌باشد (۴۶). از دیگر کاربردهای توالی‌یابی کل ژنوم، می‌توان به توالی‌یابی تک سلول در *Pre-implantation genetic diagnosis*

این نوع توالی‌یابی در سطح ترانسکریپتوم به توالی‌یابی RNA معروف است. در سال ۲۰۰۷، توالی‌یابی ژنوم پیچیده و هتروزیگوت انگور با استفاده از ترکیب روش‌های Sanger و پاپروسکوئسنینگ انجام شد. در این پروژه، برای اولین بار، توالی‌های خوانده شده با روش‌های Sanger و NGS، ژنوم پیچیده‌ی یوکاریوتی با موفقیت اسمبل شدند. در سال ۲۰۱۰، پیش‌نویسی از ژنوم *Giant panda* با استفاده از تکنولوژی NGS فراهم شد که توالی‌های هم‌پوشان اسمبل شده، حدود ۹۴ درصد ژنوم (۲/۲۵ Gb) را پوشش می‌دادند. به نظر می‌رسد که ۰/۰۵ Gb شکاف باقی مانده، مربوط به توالی‌های تکراری است. برای انجام این طرح‌ها به روش Sanger، علاوه بر هزینه‌ی بالا، زمان طولانی نیاز می‌باشد، اما با کاربرد نسل جدید روش‌های توالی‌یابی، با سرعت و هزینه‌ی کمتر، این طرح‌ها انجام می‌شود (۲۷).

تعیین توالی RNA (*RNA-seq*)

یکی از کاربردهای مهم و در حال گسترش تکنیک‌های نسل جدید توالی‌یابی، توالی‌یابی RNA کل است. با استفاده از این تکنیک، امکان بررسی تمام انواع مختلف نسخه‌های موجود در سلول اعم از RNA کد کننده و غیر کد کننده وجود دارد. تکنیک توالی‌یابی RNA برای انجام مطالعات ژنتیک و مولکولی به ویژه در موجودات غیر مدل که اطلاعات ژنتیک بسیار اندکی از آن‌ها موجود است، بسیار حایز اهمیت و پر کاربرد می‌باشد (۲۸). با استفاده از این روش، تصویر به نسبت کاملی از یک نمونه‌ی بیولوژیک حاصل می‌شود که اطلاعات قابل استخراج از آن، در برخی موارد تا حد زیادی می‌تواند جایگزین توالی‌یابی کل ژنوم گردد. این در حالی است که روش Sanger، تنها قادر به شناسایی ۶۰ درصد از نسخه‌های موجود در سلول است. به منظور توالی‌یابی RNA، پس از تبدیل این مولکول به *complementary DNA* (cDNA)، اغلب از تکنولوژی‌های توالی‌یابی ۴۵۴ و Illumina استفاده می‌شود. به علت هزینه‌ی پایین‌تر و مقدار بیشتر توالی خوانده شده توسط Illumina، این تکنولوژی به ویژه در سال‌های اخیر مورد توجه محققان مختلف قرار گرفته است (۲۹).

توالی‌یابی RNA در بیماری‌هایی مانند سرطان‌ها در زمینه‌ی تشخیص بیان یک آلل خاص یا یک رونوشت به هم متصل (*Fusion*) که به طور اختصاصی در سلول سرطانی وجود دارد، از اهمیت بالایی برخوردار است (۳۰). به عنوان مثال، اتصال ژن‌ها به هم و ایجاد رونوشت‌های ادغام شده از ویژگی‌های بسیاری از سرطان‌های خونی و سرطان‌های بافت‌های نرم می‌باشد (۳۱) که به طور معمول، تشخیص این اتصال‌های ژنی با روش‌های سیتوژنتیک با محدودیت‌هایی روبه‌رو می‌باشد، اما می‌توان با توالی‌یابی کل RNA با روش‌های NGS به تمامی این بازآرایی‌ها پی برد (۳۳-۳۲). از آن جایی که برخی از این اتصالات ژنی، در فرایندهای سرطان‌زایی نقش

(PGD) اشاره نمود (۴۷).

کاربردهای دیگر

علاوه بر کاربردهای پیش‌گفته، از نسل جدید توالی‌یابی در بررسی‌های تعامل Long non-coding RNA (lnc-RNA) و کروماتین (Chromatin isolation by RNA purification seq) یا (ChIRP-seq)، تعیین نواحی باز یا در دسترس یوکروماتین (Formaldehyde-assisted isolation of regulatory elements) یا (FAIRE .DNase-seq) و همچنین، بررسی پروفایل ژنومی موقعیت نوکلئوزوم‌ها مورد استفاده قرار می‌گیرد (۵۳-۵۱). از کاربردهای دیگر این تکنولوژی، می‌توان به حوزه‌ی متاژنومیک اشاره کرد (۵۶-۵۴).

توالی‌یابی اسیدهای نوکلئیک آزاد موجود در خون و سایر مایعات بدن
DNA و rRNA آزاد در مایعات بدن به ویژه در خون، یک رویکرد جدید تشخیصی و پیش‌آگهی دهنده در زمینه‌ی بسیاری از بیماری‌ها مانند سرطان و یا تشخیص‌های پیش از تولد در بیماری‌های تک‌ژنی و یا کروموزومی می‌باشد (۴۹-۴۸). در گذشته، با جداسازی این اسیدهای نوکلئیک، مشکل بزرگ بررسی آن‌ها در درجه‌ی اول میزان پایین آن‌ها و سپس تعقیب تنها یک سری تغییرات و جهش‌های خاص در این اسیدهای نوکلئیک آزاد بود، اما امروزه، می‌توان با استفاده از توالی‌یابی کل DNA یا rRNA موجود در این مایعات، به روش NGS به راحتی هر گونه مشکلی را تشخیص داد (۵۰-۴۹).

References

1. Avery OT, Macleod CM, McCarty M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types: Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *J Exp Med* 1944; 79(2): 137-58.
2. Watson JD, Crick FH. Molecular structure of nucleic acids; A structure for deoxyribose nucleic acid. *Nature* 1953; 171(4356): 737-8.
3. Church G, Gilbert WC. Genomic sequencing. *Proc Natl Acad Sci USA* 1984; 81(7): 1991-5.
4. Maxam AM, Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci USA* 1977; 74(2): 560-4.
5. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 1977; 74(12): 5463-7.
6. Dovich NJ. DNA sequencing by capillary electrophoresis. *Electrophoresis* 1997; 18(12-13): 2393-9.
7. Parker LT, Zakeri H, Deng Q, Spurgeon S, Kwok PY, Nickerson DA. AmpliTaq DNA polymerase, FS dye-terminator sequencing: analysis of peak height patterns. *Biotechniques* 1996; 21(4): 694-9.
8. Mardis ER. Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet* 2008; 9: 387-402.
9. Shendure J, Ji H. Next-generation DNA sequencing. *Nat Biotechnol* 2008; 26(10): 1135-45.
10. Liu L, Li Y, Li S, Hu N, He Y, Pong R, et al. Comparison of next-generation sequencing systems. *J Biomed Biotechnol* 2012; 2012: 251364.
11. Maricic T, Paabo S. Optimization of 454 sequencing library preparation from small amounts of DNA permits sequence determination of both DNA strands. *Biotechniques* 2009; 46(1): 51-7.
12. Ronaghi M, Uhlen M, Nyren P. A sequencing method based on real-time pyrophosphate. *Science* 1998; 281(5375): 363, 365.
13. Huse SM, Huber JA, Morrison HG, Sogin ML, Welch DM. Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol* 2007; 8(7): R143.
14. Ansoorge WJ. Next-generation DNA sequencing techniques. *N Biotechnol* 2009; 25(4): 195-203.
15. Liu L, Hu N, Wang B, Chen M, Wang J, Tian Z, et al. A brief utilization report on the Illumina HiSeq 2000 sequencer. *Mycology* 2011; 2(3): 169-91.
16. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, et al. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* 2012; 13: 341.
17. Daum LT, Rodriguez JD, Worthy SA, Ismail NA, Omar SV, Dreyer AW, et al. Next-generation ion torrent sequencing of drug resistance mutations in Mycobacterium tuberculosis strains. *J Clin Microbiol* 2012; 50(12): 3831-7.
18. Xuan J, Yu Y, Qing T, Guo L, Shi L. Next-generation sequencing in the clinic: promises and challenges. *Cancer Lett* 2013; 340(2): 284-95.
19. Kaiser J. A Plan to Capture Human Diversity in 1000 Genomes. *Science* 2008; 319(5862): 395.
20. Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, et al. A map of human genome variation from population-scale sequencing. *Nature* 2010; 467(7319): 1061-73.
21. Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, et al. Target-enrichment strategies for next-generation sequencing. *Nat Methods* 2010; 7(2): 111-8.
22. Zhang W, Cui H, Wong LJ. Application of next generation sequencing to molecular diagnosis of inherited diseases. *Top Curr Chem* 2014; 336: 19-45.
23. Clark MJ, Chen R, Lam HY, Karczewski KJ, Chen R, Euskirchen G, et al. Performance comparison of exome DNA sequencing technologies. *Nat Biotechnol* 2011; 29(10): 908-14.
24. Salipante SJ, Scroggins SM, Hampel HL, Turner EH, Pritchard CC. Microsatellite instability detection by next generation sequencing. *Clin Chem* 2014; 60(9): 1192-9.

25. Walsh T, Casadei S, Coats KH, Swisher E, Stray SM, Higgins J, et al. Spectrum of mutations in BRCA1, BRCA2, CHEK2, and TP53 in families at high risk of breast cancer. *JAMA* 2006; 295(12): 1379-88.
26. Bentley DR. Whole-genome re-sequencing. *Curr Opin Genet Dev* 2006; 16(6): 545-52.
27. Ossowski S, Schneeberger K, Clark RM, Lanz C, Warthmann N, Weigel D. Sequencing of natural strains of *Arabidopsis thaliana* with short reads. *Genome Res* 2008; 18(12): 2024-33.
28. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 2009; 10(1): 57-63.
29. Pease J, Sooknanan R. A rapid, directional RNA-seq library preparation workflow for Illumina[reg] sequencing. *Nat Meth* 2012; 9(3).
30. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature* 2009; 458(7239): 719-24.
31. Mitelman F, Johansson B, Mertens F. The impact of translocations and gene fusions on cancer causation. *Nat Rev Cancer* 2007; 7(4): 233-45.
32. Maher CA, Kumar-Sinha C, Cao X, Kalyana-Sundaram S, Han B, Jing X, et al. Transcriptome sequencing to detect gene fusions in cancer. *Nature* 2009; 458(7234): 97-101.
33. Pflueger D, Terry S, Sboner A, Habegger L, Esgueva R, Lin PC, et al. Discovery of non-ETS gene fusions in human prostate cancer using next-generation RNA sequencing. *Genome Res* 2011; 21(1): 56-67.
34. Steidl C, Shah SP, Woolcock BW, Rui L, Kawahara M, Farinha P, et al. MHC class II transactivator CIITA is a recurrent gene fusion partner in lymphoid cancers. *Nature* 2011; 471(7338): 377-81.
35. Furey TS. ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions. *Nat Rev Genet* 2012; 13(12): 840-52.
36. Park PJ. ChIP-seq: advantages and challenges of a maturing technology. *Nat Rev Genet* 2009; 10(10): 669-80.
37. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012; 489(7414): 57-74.
38. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 2009; 462(7271): 315-22.
39. Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res* 2005; 33(18): 5868-77.
40. Taiwo O, Wilson GA, Morris T, Seisenberger S, Reik W, Pearce D, et al. Methylome analysis using MeDIP-seq with low DNA concentrations. *Nat Protoc* 2012; 7(4): 617-36.
41. Tafe LJ. Targeted next-generation sequencing for hereditary cancer syndromes: A focus on lynch syndrome and associated endometrial cancer. *J Mol Diagn* 2015; 17(5): 472-82.
42. Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet* 2010; 11(1): 31-46.
43. Zhou X, Ren L, Meng Q, Li Y, Yu Y, Yu J. The next-generation sequencing technology and application. *Protein Cell* 2010; 1(6): 520-36.
44. Marusyk A, Polyak K. Tumor heterogeneity: causes and consequences. *Biochim Biophys Acta* 2010; 1805(1): 105-17.
45. Schubert C. Single-cell analysis: The deepest differences. *Nature* 2011; 480(7375): 133-7.
46. Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, et al. Tumour evolution inferred by single-cell sequencing. *Nature* 2011; 472(7341): 90-4.
47. Treff NR, Fedick A, Tao X, Devkota B, Taylor D, Scott RT, Jr. Evaluation of targeted next-generation sequencing-based preimplantation genetic diagnosis of monogenic disease. *Fertil Steril* 2013; 99(5): 1377-84.
48. Forsheo T, Murtaza M, Parkinson C, Gale D, Tsui DW, Kaper F, et al. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. *Sci Transl Med* 2012; 4(136): 136ra68.
49. Mosallayi M, Salehi R. Fetal cells in maternal blood: Technical and clinical aspects. *J Isfahan Med Sch* 2015; 32(313): 2165-73. [In Persian].
50. Shea JL, Diamandis EP, Hoffman B, Lo YM, Canick J, van den Boom D. A new era in prenatal diagnosis: the use of cell-free fetal DNA in maternal circulation for detection of chromosomal aneuploidies. *Clin Chem* 2013; 59(8): 1151-9.
51. Soon WW, Hariharan M, Snyder MP. High-throughput sequencing for biology and medicine. *Mol Syst Biol* 2013; 9: 640.
52. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 2007; 448(7153): 553-60.
53. Impey S, McCorkle SR, Cha-Molstad H, Dwyer JM, Yochum GS, Boss JM, et al. Defining the CREB regulon: a genome-wide analysis of transcription factor regulatory regions. *Cell* 2004; 119(7): 1041-54.
54. Huber JA, Mark Welch DB, Morrison HG, Huse SM, Neal PR, Butterfield DA, et al. Microbial population structures in the deep marine biosphere. *Science* 2007; 318(5847): 97-100.
55. Edwards RA, Rodriguez-Brito B, Wegley L, Haynes M, Breitbart M, Peterson DM, et al. Using pyrosequencing to shed light on deep mine microbial ecology. *BMC Genomics* 2006; 7: 57.
56. Barzon L, Lavezzo E, Costanzi G, Franchin E, Toppo S, Palu G. Next-generation sequencing technologies in diagnostic virology. *J Clin Virol* 2013; 58(2): 346-50.
57. Mardis ER. The impact of next-generation sequencing technology on genetics. *Trends Genet* 2008; 24(3): 133-41.
58. Johnsen JM, Nickerson DA, Reiner AP. Massively parallel sequencing: the new frontier of hematologic genomics. *Blood* 2013; 122(19): 3268-75.
59. Velasco R, Zharkikh A, Troggio M, Cartwright DA, Cestaro A, Pruss D, et al. A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. *PLoS One* 2007; 2(12): e1326.
60. Huang dW, Sherman BT, Zheng X, Yang J, Imamichi T, Stephens R, et al. Extracting biological meaning from large gene lists with DAVID. *Curr Protoc Bioinformatics* 2009; Chapter 13: Unit.
61. Li DK, Zhou Z, Miao M, He Y, Qing D, Wu T, et al. Relationship between urine bisphenol-A level and

- declining male sexual function. *J Androl* 2010; 31(5): 500-6.
62. Denver DR, Dolan PC, Wilhelm LJ, Sung W, Lucas-Lledo JI, Howe DK, et al. A genome-wide view of *Caenorhabditis elegans* base-substitution mutation processes. *Proc Natl Acad Sci U S A* 2009; 106(38): 16310-4.
63. Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, et al. Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* 2008; 452(7184): 215-9.
64. Costello AB. Getting the most from your analysis. *Pan* 2009; 12(2): 131-46.
65. Fierer N, Jackson RB. The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci U S A* 2006; 103(3): 626-31.
66. Johnsen JM, Nickerson DA, Reiner AP. Massively parallel sequencing: the new frontier of hematologic genomics. *Blood* 2013; 122(19): 3268-75.
67. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. *Nature* 2007; 449(7164): 804-10.
68. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 2010; 464(7285): 59-65.

Next-Generation Sequencing and its Applications

Meysam Mosallayi¹, Hamed Mirzaei MSc², Miganoosh Simonian¹, Majid Kheirollahi PhD³

Review Article

Abstract

DNA sequencing is an approach exploited to determine the sequence of a DNA molecule. It includes any method or technology used to identify and determine the order of the four bases of adenine, guanine, cytosine, and thymine in a strand of DNA. DNA sequencing might be used to determine the sequence of individual genes, larger genetic regions, full chromosomes, or entire genomes. Traditional sequencing methods are mainly based on the original Sanger sequencing technique which makes them very expensive and low-throughput; thus, they do not meet the needs of researchers. Consequently, with the considerable advances in molecular biology and the high demand for low-cost sequencing has encouraged the development of high-throughput sequencing (or next-generation sequencing) technologies that parallelize the sequencing process, producing thousands or millions of sequences concurrently. Next-generation sequencing enable us to rapidly sequence a large piece of DNA which could span the whole genome with the latest instruments capable of producing gigabases of data in one isolated sequencing run. Next-generation sequencing platforms have a wide variety of applications, such as whole-genome sequencing, de novo sequencing, RNA sequencing (for applications such as transcriptomics and small RNA analysis), methylation analysis, and protein-nucleic acid interaction analysis.

Keywords: Sequencing, Next-generation sequencing, 454, Solex, Sequencing by oligonucleotide ligation and detection (SOLiD)

Citation: Mosallayi M, Mirzaei H, Simonian M, Kheirollahi M. **Next-Generation Sequencing and its Applications.** J Isfahan Med Sch 2016; 33(368): 2469-80

1- MSc Student, Pediatric Inherited Disease Research Center, Research Institute for Primordial Prevention of Non-communicable Diseases AND Department of Genetics and Molecular Biology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran

2-PhD Student, Biotechnology Research Center, Mashhad University of Medical Sciences, Mashhad, Iran

3- Associate Professor, Pediatric Inherited Disease Research Center, Research Institute for Primordial Prevention of Non-communicable Diseases AND Department of Genetics and Molecular Biology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran

Corresponding Author: Majid Kheirollahi PhD, Email: mkheirollahi@med.mui.ac.ir