مقدمه‌ای بر توالی‌یابی نسل جدید و کاربردهای آن

مینا امین[۱]، الهه محمودی[۱*] و مهرداد زینلیان[۲]

[۱] کاشان، دانشگاه کاشان، دانشکده شیمی، گروه زیست شناسی سلولی و مولکولی

[۲] اصفهان، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی اصفهان، دانشکده پزشکی، گروه ژنتیک و بیولوژی مولکولی

چکیده

تعیین توالی DNA از مهم‌ترین روش‌ها در زیست‌شناسی است که توسط آن می‌توان ترتیب قرار گرفتن نوکلئوتیدها را در یک قطعه از DNA مشخص کرد. چندین روش متفاوت در تعیین توالی DNA وجود دارد که روش‌های قدیمی‌تر، مانند روش Sanger زمان‌بر و پرهزینه است و لذا محققان برای کاهش هزینه‌های توالی‌یابی و توان عملکردی بالا روش های جدیدتر توالی-یابی را ابداع کرده اند؛ امروزه توالی‌یابی را می‌توان به سه نسل تقسیم‌بندی کرد. روش های نسل جدید قادر به توالی‌یابی سریع و همزمان قطعات بزرگ‌DNA است. به کمک روش های نسل جدید می‌توان توالی‌یابی کل ژنگان، توالی‌یابی مناطق هدف، تعیین توالی از نو و سرهم‌بندی، تعیین توالی RNA و تغییرات وراژنتیکی را بررسی کرد. با توجه به اینکه هر روش از توالی‌یابی نسل جدید معایب و مزایایی دارند، از این رو مهم‌ترین امر در زمینه روش های توالی‌یابی این است که محققین بر اساس ویژگی‌های طرح پژوهشی خود، مناسب‌ترین روش توالی‌یابی را از نظر کارایی انتخاب کنند.

کلیدواژگان: تعیین توالی DNA، نسل بعدی توالی یابی، روش سَنگِر، روش ماکسام-گیلبرت، توالی‌یابی موازی بزرگ، توالی یابی تک مولکول.

* نویسنده مسئول، تلفن: ۰۳۱۵۵۹۱۳۰۴۲، پست الکترونیکی: e.mahmoodi_kh@kashanu.ac.ir

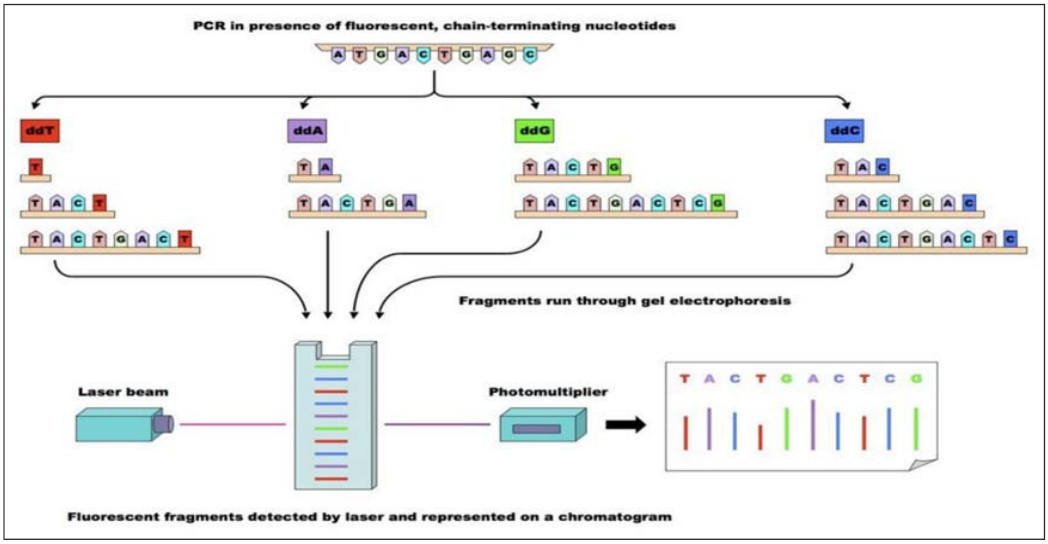
مقدمه

تعیین توالی ژنتیکی از اوائل دهه ۱۹۷۰ همواره مورد توجه دانشمندان و محققان علوم زیستی، زیست شناسی مولکولی و ژنتیک پزشکی، علوم جانوری و گیاهی بوده، روش‌های متنوعی برای این منظور ابداع شده است. توالی‌یابی نوکلئوتیدی تعیین ترتیب قرارگیری نوکلئوتیدها (A, T, C, G)در یک قطعه دزکسی‌ریبونوکلئیک‌اسید (DNA) یا ریبو-نوکلئیک‌اسید (RNA) است. روش‌های متنوع توالی یابی مستقیم و یا غیر مستقیم می توانند توالی نوکلئوتیدها را مشخص کنند. تعیین توالی نوکلئوتیدی بخش جدایی ناپذیری از علوم و تحقیقات پایه زیست شناسی است و در علوم پزشکی، ژنتیک، زیست فناوری، تشخیص‌های جنایی و پزشکی قانونی، زیست شناسی سیستمی، ویروس‌شناسی، باکتری‌شناسی و تشخیص طبی بیماری‌های ژنتیکی پیش و پس از تولد کاربرد دارد.

پیشرفت چشمگیر و بسیار سریع در تعیین توالی که با تکنولوژی مدرن به دست آمد، زمینه یافتن توالی کامل (ژنگان) موجودات گوناگون شامل ژنگان انسان و سایر جانوران، گیاهان و گونه‌های ویروسی و باکتری را فراهم کرد. روش‌های قدیمی توالی‌یابی بسیار پیچیده و وقت‌گیر بوده‌اند و در طول زمان از شروع آن در اوائل دهه ۱۹۷۰ تاکنون دائماً تغییر یافته‌اند و روشهای نسل جدید توالی یابی (NGS= new generation sequencing) توالی ها بسیار ساده‌تر و به مراتب سریع‌تر و دقیق‌تر انجام پذیرند. بر پایهٔ پیشرفت ها سه نسل توالی یابی معرفی شده است: نسل اول توالی‌یابی از ۱۹۷۷ تا ۲۰۰۷ (روش سنگر، ماکسام – گیلبرت)، نسل دوم توالی‌یابی از ۲۰۰۷ تا ۲۰۱۳ (Massive parallel sequencing) و نسل سوم توالی‌یابی از ۲۰۱۲ تا کنون، روش تک مولکول (Single molecule sequencing) است. یکی از نخستین نمونه‌های تعیینِ توالی کامل، تعیین توالی RNA مربوط به ژنگان باکتریوفاژMS2 بود که در دانشگاه GHENT بلژیک بین سال‌های ۱۹۷۲ و ۱۹۷۶ انجام شد، این موضوع نقطه عطف توالی‌یابی ژن‌ها بود(۱).

نسل اول توالی‌یابی

اولین نسل از تکنولوژی توالی‌یابی توسط سنگر و ماکسام - گیلبرت بنیان نهاده شد که به ترتیب بر اساس سنتز توالی و ایجاد برش در توالی بودند. تا قبل از سال ۲۰۰۸ روش سنگر در پژوهش‌های زیستی روش غالب بود.



شکل ۱- توالی‌یابی بخشی از DNA توسط روش سنگر: نمونه در بردارنده DNA تک رشته را با تقسیم به ۴ بخش یکسان و به درون چهار لوله آزمایش مختلف انتقال داده، سپس همه‌ی مواد مورد نیاز برای همانندسازی نیز به درون هر لوله اضافه می‌شود. این قطعات بر روی ژل پلی‌آکریل‌آمید سوار می‌شوند(۱).

چهار برچسب رنگی استاندارد فلوئورسنت که هر رنگ آن به یکی از ۴ باز شرکت‌کننده در ساختمان DNA بود با استفاده ازالکتروفورز مویرگی به عنوان روشی برای شناسایی اتوماتیک بازها به صورت تجاری توسط شرکتهای Applied Biosystems، Life Technologies و Beckman Coulter عرضه می‌شود. کاربرد دی‌دئوکسی نوکلئوزیدتری‌فسفات در روش سنگر به عنوان ختم دهنده زنجیر DNA است. روش کلاسیک ختم رشتهٔ در حال سنتز نیاز به DNA تک رشته‌ای همچون الگو، آغازگر و DNA پلیمراز، دی‌دئوکسی‌نوکلئوزیدتری‌فسفات و نوکلئوتیدهای تغییر یافته (dideoxyNTPs) برای ختم ساخته شدن رشته DNA دارد. (شکل ۱) بنیان این روش، به کارگیری نوکلئوتیدهای ویژه ای به نام ddNTPs (dideoxynucleotids) بوده که فاقد گروه OH بر روی کربن ′ ۳ خود هستند و اگر در ساختار یک اسید نوکلئیک قرار گیرند این زنجیره دیگر بلندتر نخواهد شد (۱ و ۲).

توالی‌یابی بر اساس روش ماکسام-گیلبرت، بر پایه هضم شیمیایی بوده و نیازی به پلیمریزاسیون قطعه مورد نظر ندارد و یک فسفات رادیواکتیو به انتهای 5′ از DNA دو رشته اضافه می شود (شکل ۲) (۳, ۴)؛ نخستین پروژه‌ی توالی‌یابی کامل ژنگان انسان (ژنگان کریگ ونتر) در سال ۲۰۰۷ به وسیله‌ی روش توالی‌یابی سنگر به اتمام رسید.

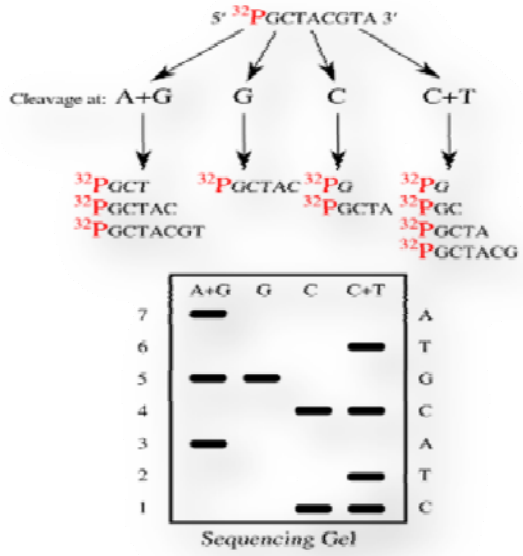
گرچه اکنون دیگر روش غالب توالی‌یابی بر روی بسترهای NGS انجام می‌شود اما هنوز به خاطر صرفه جویی در هزینه ها در پروژه‌هایی که هدف توالی‌یابی یک توالی مشخص از DNA است مانند تشخیص بیومارکرها، آنالیز مسیرها، کاربردهای تشخیصی بالینی، از بستر CE سنگر استفاده می‌شود.

نسل دوم توالی یابی

این نسل شامل چهار روش به شرح زیر است:

روش pyrosequencing، روش Semiconductor Sequencing (Ion Torrent) ، روش illumine و روش ABI (Applied Biosystems) .

نسل دوم توالی‌یابی توسط دستگاه Roche 454 و تحت عنوان پایروسیکوئنسینگ به بازار معرفی شد. Roche 454 از امولسیون PCR جهت تکثیر کلون وار توالی هدف استفاده می‌کند. ماشین توالی‌یابی شامل تعداد فراوانی از چاهک-هایی با حجمی در مقیاس پیکولیتر بود که هر کدام از این چاهک ها در درون خود دارای یک گویچه کوچک و آنزیم‌های توالی‌یابی هستند (شکل۳). پایروسکوئنسینگ از آنزیم لوسیفراز که نور تولید می‌کند جهت تشخیص بازهای تازه سنتز شده در DNA استفاده می‌کند (۵ و۶).

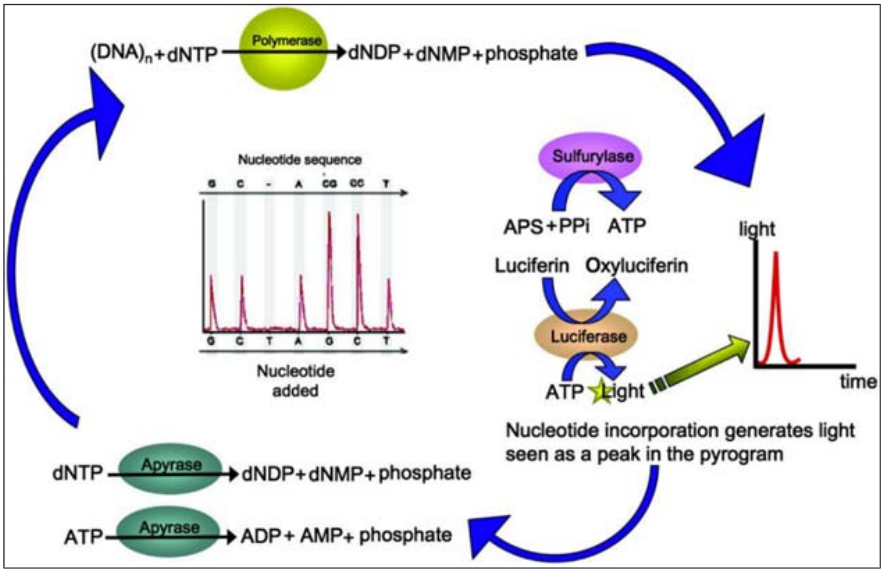


شکل۲- توالی‌یابی به روش ماکسام-گیلبرت: ابتدا فسفات '5 یا '3 انتهایی نشاندار می‌شود، سپس به وسیله‌ی اندونوکلئازهای برشی، DNA به دو قطعه ی نامساوی تبدیل می‌شود که هر یک دارای یک فسفر نشاندار است. در مرحله بعد، دو رشته از هم جدا می‌شوند و قطعات توسط الکتروفورز از هم جدا می شوند. در نهایت به روش شیمیایی خاصی نوکلئوتیدها را جدا و توالی DNA را تعیین می‌کنند(۲).

توالی‌یابی به وسیله‌ی ایلومینا از طریق خاتمه‌دهنده برگشت پذیر است، ایلومینا از تکثیر خوشه‌ای توالی هدف بر روی یک سطح جامد استفاده می‌کند (شکل ۴، الف و ب). توالی‌یابی با استفاده از افزودن چهار نوع نوکلئوتید که هرکدام با یک رنگ فلوئورسنت ویژه نشانه گذاری شده، دارای بازهای خاتمه دهنده قابل بازگشت هستند (شکل ۴، ج). برخلاف روش پایروسکوئنسینگ در روش ایلومینا سنتز DNA در هر مرحله تنها با افزودن یک نوکلئوتید صورت می‌گیرد. بعد از ثبت تصاویری که بر اثر افزوده شدن بازهای فلوئورسنت به DNA ایجاد شده اند، فلوئوروفور به همراه خاتمه دهنده‌های قابل بازگشت به کمک مواد شیمیایی از DNA شسته شده، شرایط برای آغاز دور بعدی مهیا می‌شود (۹،۱۰).

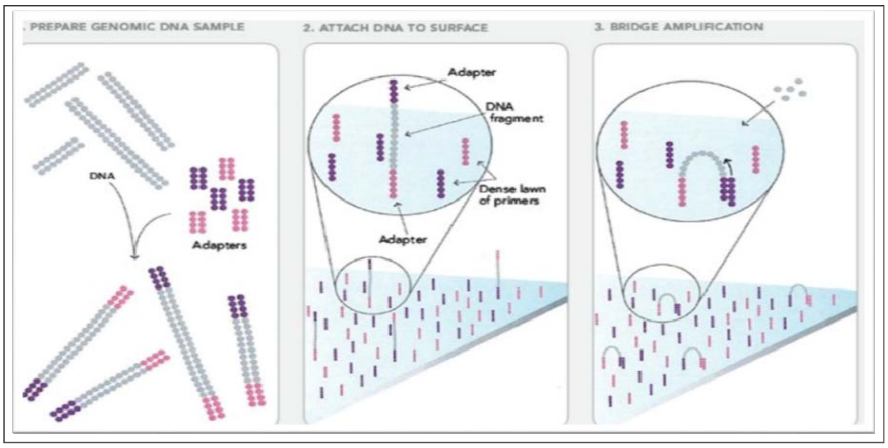
شرکت های Applied Biosystem/Life Technologies' SOLiD technology واکنش اتصال برای توالی یابی را با استفاده از پروب‌هایی به کار می گیرند که از همه‌ی اولیگونوکلئوتیدهای محتمل مربوط به یک توالی با طول ثابت، بر اساس محل توالی نشانه‌گذاری شده‌اند (شکل ۵). اولیگو نوکلئوتیدها بعد از قرار گرفتن درکنار یکدیگر به هم متصل می‌شوند. اتصال دلخواه قطعات به وسیله لیگاز منجر به ثبت موقعیت توالی‌ها می‌شود. DNA به صورت کلون وار بر روی گوی‌های فلزی کوچک درون امولسیون PCR تکثیر می‌یابد به نحوی که هر گویچه فقط نسخه های یک قطعه از DNA را دارند. سپس گویچه ها روی اسلایدهای شیشه ای قرار می‌گیرند و توالی‌شان تعیین می شوند. توالی ها از نظر طول و کمیت توالی یابی با روش ایلومینا قابل مقایسه هستند(۱۱، ۱۲).

روش Semiconductor Sequencing (Ion Torrent)، از تغییرات pH ، رها شدن یون H^+ به هنگام سنتز رشته تازه در محیط، برای شناسایی نوکلئوتیدها استفاده می‌شود و در پایان مهره های فلزی دربرگیرنده بیش از یک میلیون DNA تک رشته‌ای به درون ریزچاهک‌هایی انتقال می‌یابد؛ در زیر این چاهک‌ها یک حسگر فوق حساس یونی تعبیه شده است که سریعاً تغییرات pH را ثبت می‌کند و در پایان این تغییرات به صورت یک پیک در نمایشگر آشکار می-شود. چون در این روش نیز در هر مرحله نوکلئوتیدهای همانند و شناخته شده به مخلوط واکنش افزوده می‌شود، می‌توان بر اساس تولید این پیک، توالی DNA را هنگام سنتز به دست آورد (شکل۶)(۱، ۷ و ۸).

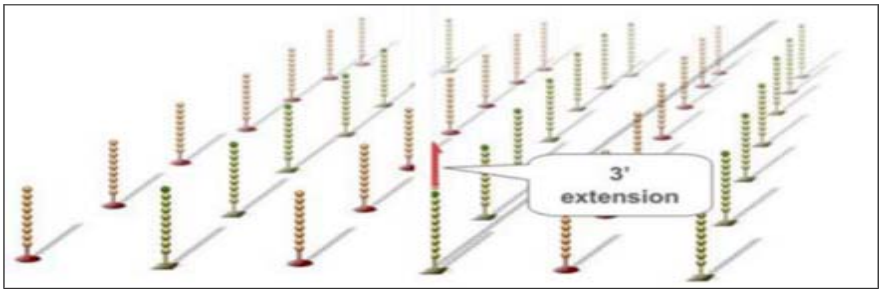


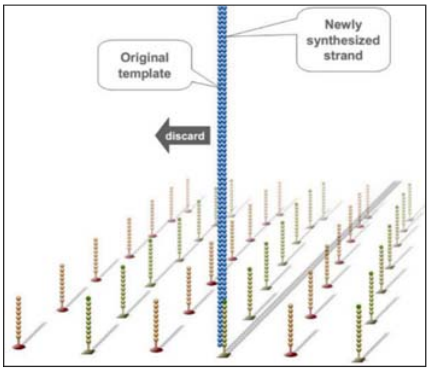
شکل ۳- پایروسکوئنسینگ: این روش براساس ساطع شدن نور استوار است. به طوری که با وارد شدن نوکلئوتید همسان به رشته روبرو گروه پیروفسفات آزاد شده به همراه آدنوزین فسفوسولفات (APS) موجود در مخلوط واکنش در تولید ATP استفاده می شود که این سیگنال نوری به صورت یک پیک ثبت می‌شود که بر پایهٔ آن، توالی DNA در هنگام سنتز به دست می‌آید، توالی رشته الگو براساس نوکلئوتید اضافه شده تعیین می‌شود (۹).

الف

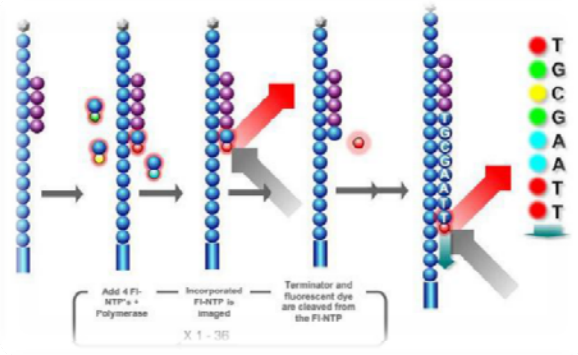


ب





ج



شکل ۴- مراحل روش illumine

الف: تکنیک illumine: قبل از تک رشته ای شدن، به دو سمت قطعات DNA آداپتور متصل می‌شود. سپس این قطعات تک رشته ای شده، به صورت تصادفی بر روی یک سطح جامد توزیع می‌شوند که این سطح جامد به طور متراکم با توالی‌های مکمل آداپتور پوشیده شده است؛ ب: تکثیر خوشه‌ای توالی هدف بر روی یک سطح جامد، پرایمرها از طرف گروه فسفات '۵ به تراشه متصل شده اند لذا ناحیه هیدروکسی '۳ آن آزاد است؛ ج: آغاز عمل پلیمراز و رها شدن فلوروفرهای متصل به نوکلئوتید و ردیابی نور آزاد شده

انواع روش‌های مربوط به نسل دوم توالی‌یابی از نظر حجم کارایی، طول خوانش و هزینه از یکدیگر متفاوت‌اند. این روش‌ها عموماً از نظر حجم کاری پربازده تلقی می‌شوند اما قیمت دستگاه‌های آن‌ها بسیار گران و در محدودهٔ نیم تا یک میلیون دلار است. ثبت و خوانش سیگنال‌ها در هر دو روشی که از نشانگرهای فلوئورسنت یا تجزیه پیرو فسفات استفاده می‌کنند نیازمند ردیاب های نوری است. بسترهای نسل دوم توالی‌یابی گرچه دربسیاری از کاربردهای پژوهشی موفق بوده اند اما ردیاب های هزینه بالا، به خاطر سختی آماده سازی نمونه ها و مواد شیمیایی (نشانگذاری فلوئورسنت و واکنش‌های آنزیم سوبسترایی) پیچیدگی کار با ردیاب های نوری و محدودیت طول خوانش دچار محدودیت هایی هستند.

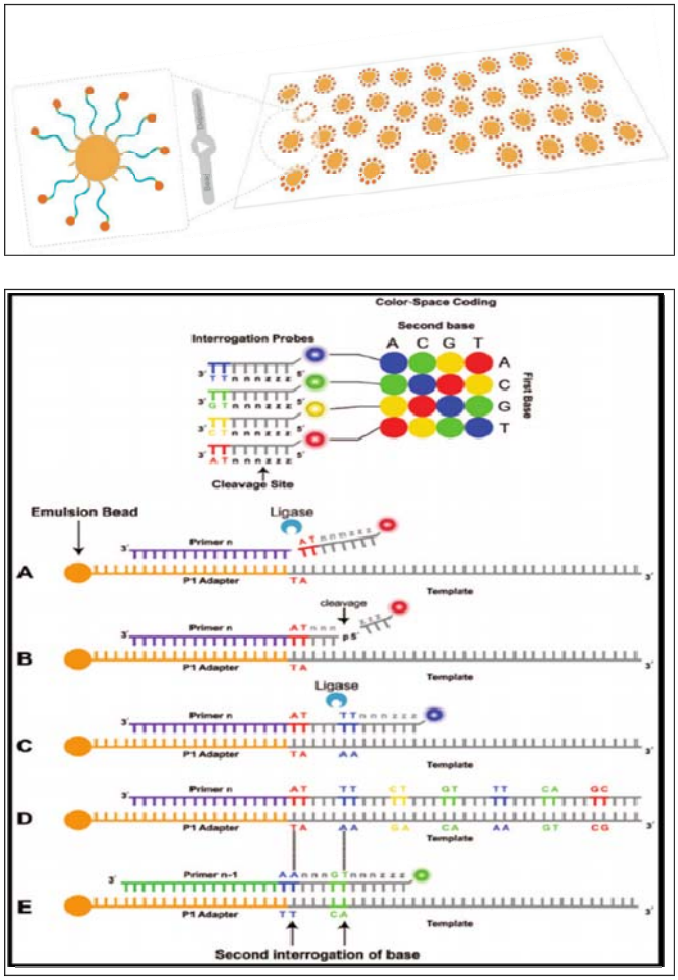
نسل سوم تکنولوژی توالی‌یابی

گرچه روش‌ها و ابزارهای نسل دوم بسیار ارزان و بازدهی بالا دارند ولی در تمامی مراحل باید پیش از ورود به مرحله توالی‌یابی تعداد رشته ها به اندازه ای افزایش یابد که سیگنال برامده قابل ردیابی باشد. این موضوع علاوه بر افزودن به زمان کل فرآیند باعث خطا در توالی‌یابی می-شود، بنابراین حذف این روش باعث مرتفع شدن این مشکلات خواهد شد. در نسل سوم که توالی‌یابی تک مولکول نامیده شده است، نیازی به افزایش شمار نسخه ها پیش از مرحله توالی‌یابی نیست و مستقیماً پس از آماده-سازی نمونه ها مرحله توالی‌یابی آغاز می‌شود بنابراین اندازه خطا بسیار پایین است(۱۳) .

این نسل شامل روش: (Heliscope (Helicos Single molecule sequencing devi ، روش (SMRT (Single Molecule Real Time یا Pac Bio sequencing و روش Nanopore Sequencing است.

(PacBioRS) یک روش توالی‌یابی تک پلیمراز- تک ملکول به صورت Real time است که می‌تواند 1000 bp را بخواند. هر تراشه دارای یک حالت صفر موجی است (ZMW)[1] که نانوساختارها را به منافذ ۱۰۰ نانومتری هدایت می‌کند که درون آنها DNA پلیمراز سنتز توالی نوکلئوتیدهای متصل به فسفر نشان دار شده با فلوئوفور انجام می‌دهد. علاوه بر مشخص شدن توالی DNA، نشان دادن سرعت اتصال نوکلئوتید ها ممکن است در آینده به اطلاعات اپی‌ژنتیکی (مانند الگوی متیلاسیون) رشته های DNA کمک کند. این قالب توانایی توالی‌یابی mRNA را نیز داراست که این کار با جایگزین کردن ریبوزوم به جای DNA پلیمراز صورت می‌گیرد (شکل ۷)؛ چنین وسیله‌ای با چنین ساختاری بسیار گران خواهد بود (۱،۵،۱۳) .

[1] zero-modo waveguide



شکل ۵- روش ABI (Applied Biosystems): این روش بر پایه پیوستن پروب‌های الیگونوکلئوتیدی نشاندار شده با ۴ رنگ مختلف فلوئورسانس (قرمز، آبی، نارنجی و سبز) استوار است. هر پروب جایگاه دو باز را در زنجیره مشخص می‌کند. هر کدام از این دانه‌ها که حاوی یک قطعه DNA تک رشته این است، در درون یک قطره آب در امولسیون روغن قرار می‌گیرند و واکنش PCR جهت تکثیر قطعه DNA درون این میکرورآکتور انجام می‌شود (۱۲).

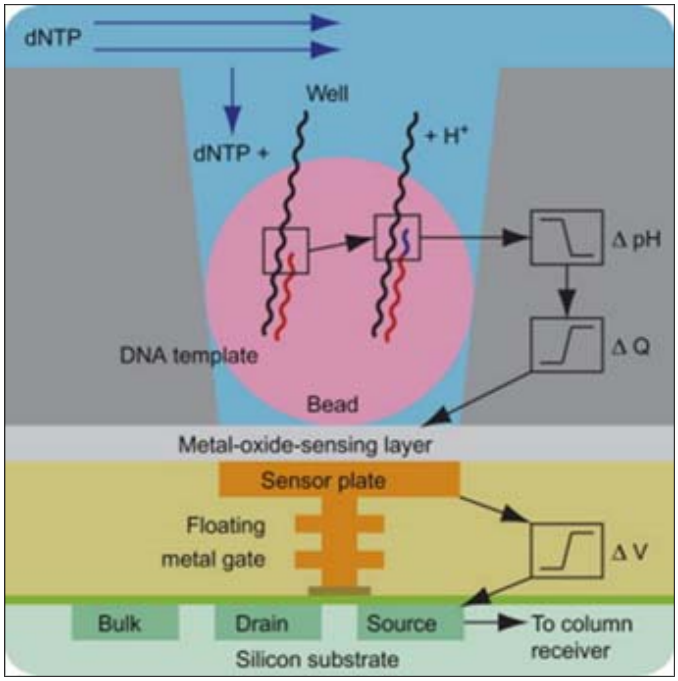
ژنومیکس کامل، یک روش ترکیبی از توالی‌یابی با بالاترین توان ورودی را در میان سومین نسل توالی‌یابی از خود نشان داده است. در این روش از چرخه های تکثیر توالی-های کوچک DNA در جایی که نانوتوپ (nanoball) نامیده می‌شود، استفاده می‌شود. این روش اجازه می‌دهد که تعداد زیادی از نانوتوپ‌های DNA در هر مرحله توالی‌یابی شوند که این کار با هزینه کمی صورت می‌گیرد. این روش به طور موفقیت آمیزی در کاربردهای بالینی توالی‌یابی ژنگان نظیر توالی‌یابی ژنگان کامل مورد استفاده قرار می‌گیرد. با این حال، خواندن نقشه‌های توالی‌های کوتاه در پایگاه داده ژنومیک رفرنس به خصوص در آنالیز DNA تومور می-تواند مشکل باشد. شاید پرکاربردترین و کم هزینه‌ترین روش در تشکیل ماشین ژنگانی شخصی[1] (PGM) یک ابزار ضروری در آزمایشگاه‌های تحقیقاتی و بالینی آینده باشد (۱).

شیمی توالی‌یابی در حقیقت با تکنولوژی جریان یون و آزاد شدن پروتون در طی اتصال هر نوکلئوتید توسط DNA پلیمراز مرتبط می‌باشد. در روش ریزآرایه[2] میکرو چاهک

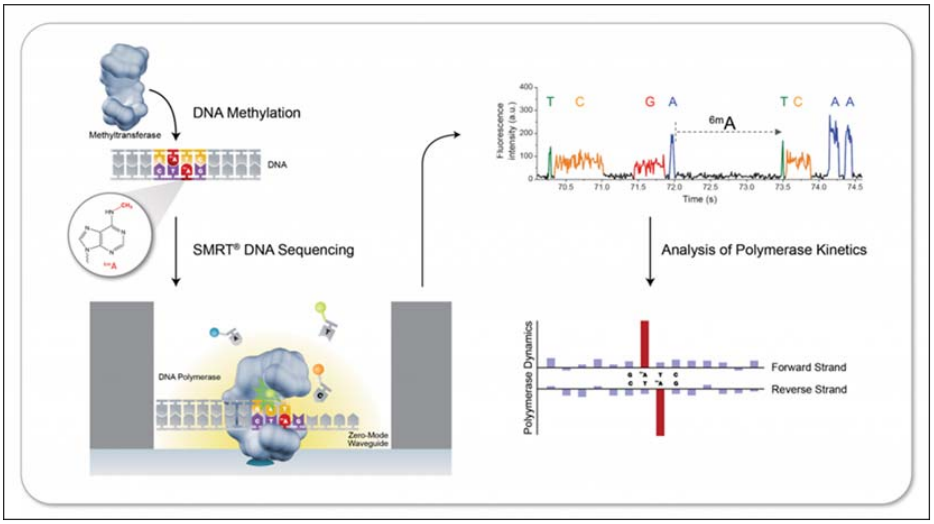
[1] Personal genomic machine
[2] .Microarray

های انفرادی به DNA پلیمراز اجازه می‌دهد تا به صورت توده ای قطعات DNA هدف را تکثیر کند. در زیر هر میکرو چاهک ترانزیستور حساس به جریان یونی (ISFET)تغییرات pH ناشی از آزاد سازی پروتون را شناسایی کرده و تغییر پتانسیل (DV) با اندازه گیری مستقیم وقایع اتصال نوکلئوتید ثبت می‌شود (شکل ۸)؛ این سیستم نیازی به نشانه‌گذاری نوکلئوتید ها ندارد و هیچ شناساگر نوری مورد نیاز نیست. گنجایش توالی‌یابی PGM جریان یونی برای کمتر از K100 برای پروژه‌های تحقیقی با مقیاس کم یا ازمایشگاه‌های تشخیص بالینی کافی می‌باشد (۱۳،۱۴ .) ورود آن به بازار متکی به شروع NGS به عنوان کالایی با کاربردهای بالینی و زیست پزشکی می‌باشد. به عنوان یک ویژگی مشترک در سایر سیستم ها، آن اداپتورها بارکدی چند گانه دارند که اجازه تست‌های پویای چندین نمونه را می‌دهد. سایزهای تراشه‌های موجود (۳۱۴–۳۱۸) و 10-1000 MB اطلاعات توالی در هر ران می‌باشد. از آنجایی که عملیات جریان یونی موجود یک فرایند فشرده است اخیراً اتوماسیون با آماده‌سازی کتابخانه (کیت آماده-سازی توالی‌یابی یک مرحله‌ای کتابخانه) جهت تسهیل فرایند در دسترس قرار گرفته است. محدودیت‌ها شامل طول خواندن کوتاه (100-200 bp) و مشکلات فنی در خواندن از طریق توالی‌های تکراری زیاد و هوموپلیمرها می‌باشد که اخیراً در بهبود بخشیدن آن تحقیقاتی صورت گرفته است (۱).

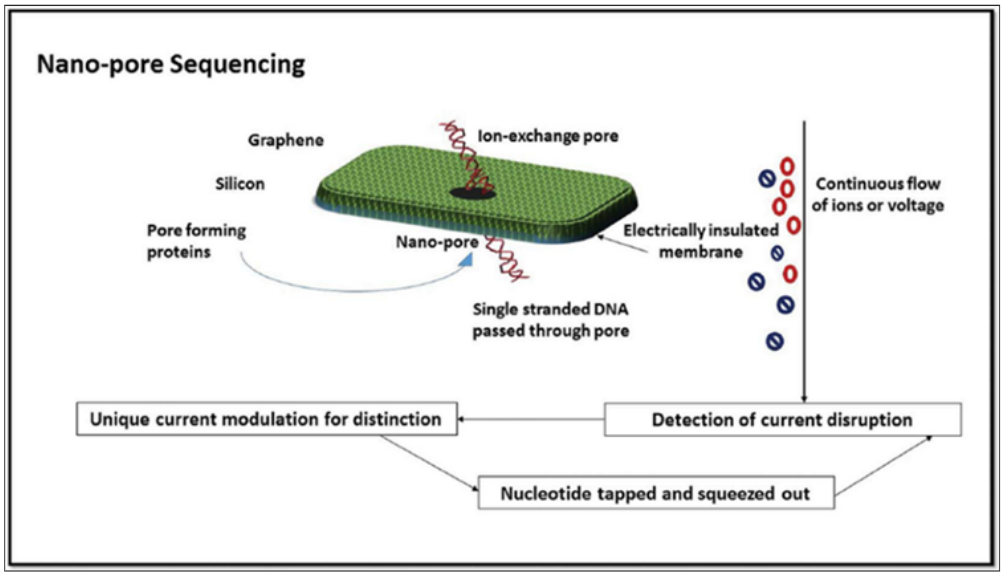
تکنیک Heliscope (Helicos Single molecule sequencing devi) : این تکنولوژی از توالی‌یابی تک مولکول واقعی بهره می‌برد. قطعات DNA به همراه دم پلی A بر روی سطح یک محبس ثابت می‌گردند و به دنبال آن توالی‌یابی بر اساس افزایش طول قطعه به کمک شستشو های چرخه ای با استفاده از نوکلئوتیدهای نشاندار شده با فلوئورسنت صورت میگیرد (شکل ۹)؛ اگرچه خوانش‌ها کوتاه هستند پیشرفت‌های اخیر در این روش منجر به افزایش دقت خوانش از طریق هوموپلیمرها شده است و همچنین امکان توالی‌یابی RNA را نیز فراهم ساخته است (۵ و ۱۳)



شکل ۶- روش Semiconductor Sequencing (Ion Torrent): در این روش، از تغییرات pH به هنگام سنتز رشته تازه در محیط برای توالی یابی استفاده می‌شود و در پایان مهره های فلزی در برگیرنده بیش از یک میلیون DNA تک رشته به درون ریزچاهک‌هایی انتقال می‌یابد که در زیر این چاهک‌ها یک حسگر فوق حساس یونی تعبیه شده است و در پایان این تغییرات pH به صورت یک پیک در نمایشگر آشکار می‌شود (۱).



شکل ۷- تکنیک SMRT (Single Molecule Real Time) یا Pac Bio sequencing : تعیین توالی همزمان با سنتز، مخلوطی از نوکلئوتیدها که هر کدام با رنگ خاصی نشاندار شده‌اند به همراه سایر عوامل مورد نیاز برای پلیمریزاسیون جهت تعیین توالی به داخل هر یک از این چاهک‌ها افزوده می‌شوند(۱۳.)



شکل ۸- روش Nanopore Sequencing، از یک غشاء بسیار نازک که حاوی کانال‌هایی به قطر ۱ الی ۲/۵ نانومتر می باشد بهره می برد که در دو سمت این کانال‌ها، ۲ الکترود تعبیه شده است. این امر منجر به ایجاد یک جریان یونی در درون این کانال‌ها می‌شود. زمانی که DNA به صورت تک رشته‌ای از درون این سوراخ‌های بسیار ریز عبور می‌نماید، بر پایه میزان مقاومت هر یک از نوکلئوتیدها، جریان یونی تغییر می‌یابد و این تغییرات توسط دستگاهی ثبت می‌گردد (۱۴).

کاربردهای متنوع NGS

تکنیک NGS کاربردهای متنوعی را مقدور می‌سازد و به محققان اجازه می‌دهد تا هر نوع موضوعی را در ارتباط با ژنگان، ترانسکریپتوم و اپی‌ژنگان مرتبط با هر موجودی مورد سوال قرار دهند. برنامه تعیین توالی تا حد زیادی از طریق آماده‌سازی کتابخانه‌های توالی و از طریق آنالیز داده دیکته شده‌است که اساساً مرحله تعیین توالی واقعی بدون تغییر باقی‌مانده است. تعدادی کیت آماده‌سازی کتابخانه استاندارد وجود دارد که پروتکل‌هایی را برای تعیین توالی کل ژنگان، mRNA، نواحی مورد هدف از قبیل اگزوم‌ها،

نواحی انتخاب شده اختیاری، نواحی متصل به پروتئین و غیره پیشنهاد می‌کنند. بسیاری از محققان برای رسیدن به اهداف خاص، پروتکل‌های جدیدی را برای جداسازی نواحی خاصی از ژنگان را که مرتبط با یک عملکرد زیست‌شناختی خاص است را توسعه داده‌اند.

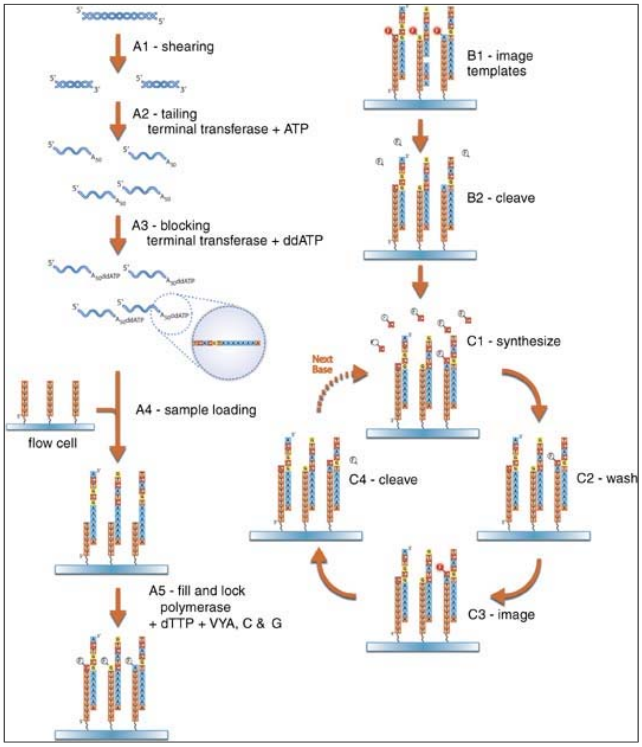
کاربردهای متنوع NGS

۱) توالی‌یابی ژنگان کامل

در این روش کل توالی ژنگان یک موجود (ژنگان هسته‌ای به همراه DNA میتوکندریایی در سلول‌های جانوری و DNA کلروپلاست در سلول‌های گیاهی) تعیین می‌شود. نخست دست یافتن به واریانت های ژنتیکی که افراد بیمار را نسبت به دیگر افراد جامعه به یک بیماری حساس می‌کند، به عنوان یک مشکل مورد توجه بوده است، با دست‌یابی به روش های NGS و توالی‌یابی کل ژنگان از راه آن توانایی بررسی واریانت-هایی با فراونی پایین‌تر و بر روی تعداد افراد بیشتری فراهم شده است(۱,۱۵)

۲) توالی‌یابی مناطق هدف

باید توجه داشت که توالی‌یابی NGS روش بهینه و فراگیرتری برای آنالیز ژنگان است، اما این روش در بسیاری از نقاط در دسترس نیست. بنابراین توالی‌یابی یک سری از توالی‌های خاص که جهش یا تغییرات دیگر در آنها موجب بیماری می‌شود مورد توجه قرار گرفته است، در این روش به جای توالی‌یابی کل ژنگان تنها اگزون‌های نمونه (معمولاً انسان) توالی-یابی می‌شوند. اهمیت این روش عمدتاً به دلیل ارزان بودن آن نسبت به توالی‌یابی کل ژنگان می‌باشد. مجموعه تمامی اگزون‌های ژنگان انسان تنها معادل ۱ درصد کل ژنگان آن می‌باشد؛ از طرفی بیش از ۸۰ درصد از جهش‌های بیماری‌زا در این نقاط اتفاق می-افتد. لذا در بسیاری از بیماری‌های ژنتیکی بهتر است به جای whole genome sequencing از exome sequencing استفاده شود.



شکل ۹- روش Heliscope، با به کارگیری آنزیم ترمینال ترنسفراز به انتهای ′۳ این قطعات تک رشته، دزوکسی‌آدنوزین‌تری‌فسفات را اضافه کرده تا یک دم پلی dA ایجاد شود. جهت تعیین توالی، نوکلئوتیدهای ختم دهنده برگشت‌پذیر که همگی توسط یک رنگ فلوئورسانس نشاندار شده‌اند به همراه آنزیم پلیمراز به این صفحات افزوده می‌شوند تا فرآیند تعیین توالی بر پایه تکثیر آغاز گردد (۱۳).

امروزه از فرم های گوناگون NGS برای Exome sequensing در بیماری‌هایی که چندین ژن ناهمسان مستعد برای جهش دارند و یا بدون نقاط Hot spot برای جهش هستند، استفاده می‌شود که می توان به مواردی مانند بیماری‌های نقص ایمنی و سرطان‌هایی مانند کلورکتال و یا سینه اشاره کرد (۱،۱۶).

۳) تعیین توالی از نو و سرهم بندی[1]

توالی‌یابی از نو، به معنای تولید نخستین توالی ژنومی یا ترانسکریپتومی برای موجوداتی که داده‌ی ژنتیکی اندکی از آنها در اختیار داریم می باشد، این نوع توالی‌یابی به عنوان توالی‌یابی RNA شناخته شده است که در سال ۲۰۰۷ توالی‌یابی ژنگان پیچیده و هتروزیگوت انگور با به‌کارگیری ترکیب روش‌های Sanger و پایروسکوئنسینگ انجام شده است و برای نخستین بار توالی‌های خوانده شده با روش‌های Sanger و NGS از ژنگان پیچیده‌ی یوکاریوتی با موفقیت سرهم بندی شدند(۱,۱۷)

۴) تعیین توالی RNA

یکی از کاربردهای مهم NGS، توالی‌یابی RNA کل است؛ با به‌کارگیری این روش امکان بررسی تمام انواع مختلف نسخه های موجود در سلول اعم از RNAرمزگذار و غیررمزگذار وجود دارد، به منظور توالی‌یابی RNA پس از تبدیل این مولکول به cDNA بیشتر از تکنولوژی‌های توالی‌یابی ۴۵۴ و Illumina استفاده می‌شود. با استفاده از RNA-Seq می‌توان در زمان بسیار اندک به پروفایل بیانی کلیه ژن‌های ارگانیسم به صورت کمی و با هزینه اندک دست یافت (ترنسکریپتوم). در این روش تنها به RNA کل نیاز است و بقیه مراحل آزمایش توسط کمپانی انجام می‌شود. این روش مزایای بسیاری نسبت به روش منسوخ Microarray دارد. در این روش علاوه بر اندازه‌گیری کمی بیان ژن‌ها، تعیین توالی نیز صورت می‌گیرد. بر خلاف real-time PCR نیازی به طراحی پروب و یا پرایمر نیست. دقت این نوع ترنسکریپتومیکس به مراتب بیشتر از Microarray است. و از همه مهم‌تر هزینه آن بسیار مقرون به صرفه‌تر است. حتی در مقایسه با real time PCR اگر تعداد ژن‌های کاندید حدوداً بیش از ۱۰ عدد باشد روش RNA-Seq ارزانتر (و البته سریعتر، بسیار ساده تر و حجم داده‌ی دریافتی بسیار بیشتر) خواهد بود. توالی‌یابی RNA در بیماری‌هایی مانند سرطان‌ها در زمینه‌ی شناسایی بیان یک آلل خاص یا یک رونوشت به هم متصل که به طور اختصاصی در سلول سرطانی وجود دارد، از اهمیت بالایی برخوردار است و به عنوان مثال اتصال ژن‌ها به هم و ایجاد رونوشت‌های ادغام شده از ویژگی‌های بسیاری از سرطان‌های خونی و سرطان‌های بافت‌های نرم است و از این رونوشت‌ها می‌توان به عنوان نشانگرهای تشخیصی و کمک کننده به تعیین نوع درمان نیز استفاده نمود که با توجه به کارامدی NGS در تعیین توالی RNA کل میتوان از آن به عنوان یک ابزار مناسب در تشخیص و کمک کننده در درمان یاد کرد (۱،۱۸).

۵) بررسی تغییرات اپی ژنتیک

ارتباط بین پروتئین و DNA یک برهمکنش بیولوژیک است که نقش عمده ای در قابل دسترس بودن DNA و تنظیم بیان ژن‌ها دارد. برای بررسی پروفایل پروتئین‌های متصل شده به DNA در مقیاس ژنومی و همچنین تغییرات پروتئین‌های هیستونی و نوکلئوزوم‌ها، از تکنیک combining chromatin immunoprecipitatiom استفاده می‌شود. به این ترتیب که پس از رسوب‌دهی ایمونوکروماتین و تفکیک پروتئین‌ها، DNA به کمک روش NGS توالی‌یابی می‌شود (۱،۱۹).

۶) توالی‌یابی اسیدهای نوکلئیک آزاد در خون و سایر مایعات بدن

بررسی وجود DNA و RNA آزاد در مایعات بدن به ویژه در خون، یک رویکرد جدید تشخیصی و پیش آگهی دهنده در زمینه‌ی بسیاری از بیماری‌ها مانند سرطان و یا تشخیص‌های پیش از تولد در بیماری‌های تک ژنی و یا کروموزومی است. در گذشته با جداسازی این اسیدهای نوکلئیک، مشکل بزرگ بررسی آن‌ها در درجه‌ی اول میزان پایین آنها و سپس

[1]. Denovo sequencing and assembly

نشان‌دار کردن آنها تنها با یک سری از تغییرات و جهش‌های خاص در این اسیدهای نوکلئیک آزاد بود، اما امروزه می‌توان با استفاده از توالی‌یابی کل DNA یا RNA موجود در این مایعات به روش NGS به راحتی هر گونه مشکلی را تشخیص داد(۲۰).

۷) آنالیز جمعیت میکروبی

با استفاده از NGS، متاژنومیکس که به مطالعه جمعیت میکروبی گفته می‌شود بسیار سریعتر و عملی‌تر شده است. در این روش کل DNA از نمونه محیطی (آب، خاک، سطح بدن انسان و موجودات و غیره) استخراج و بقیه مراحل توسط کمپانی جهت تعیین توالی کل ژنگان‌های موجود (متاژنگان) انجام می‌شود. یک روش مناسب و ایده آل برای تعیین توالی ژنگان‌های ویروسی، باکتریایی و مخمر است و در حال حاضر مقادیر زیادی از داده‌های در دسترس از NGS به کشف جدید اتصال رونوشت‌ها، انواع جهش‌ها و تعاملات پاتوژنی منجر شده است. برای مثال NGS موفق به شناسایی ژنگان *E. coli* O104:H4 شده است. همچنین NGS قادر است کل ژنگان و ارتباطات فنوتیپی را در بین ارگانیسم‌های نزدیک و مکانیسم‌های ژنتیکی را بررسی کند. این روش جدید متاژنومیکس امکان شناسایی ویروس‌های غیر منتظره بیماری زا و ویروس‌های جدید انسانی را دارد(۱).

بحث و نتیجه گیری

اختراع NGS دانشمندان را قادر ساخت تا سیستم‌های بیولوژیکی را در سطوحی مطالعه کنند که هرگز امکان پذیر نبود. با پیشرفت تکنولوژی، تعدادی از روش‌های ابتکاری آماده‌سازی نمونه و الگوریتم‌های آنالیز داده، دامنه وسیعی از کاربردهای علمی را فراهم ساخت تا محققان در تعدادی از حوزه های بیولوژیک که تاکنون پاسخی به آنها داده نشده بود به کشفیات جذابی دست پیدا کنند. داده‌های خروجی NGS با نرخی بیش از ۲ برابر نسبت به زمان اختراع آن در حال افزایش بوده است. در سال ۲۰۰۲، یک توالی‌یابی واحد می‌توانست حداکثر حدود ۱ گیگا باز (Gb) داده تولید کند اما در سال ۲۰۱۱ این نرخ به حدود یک ترا باز (Tb) داده در یک توالی‌یابی رسیده است.

NGS با این قابلیت حجم بالایی از داده‌های توالی را تولید می‌کند و محققان را قادر می‌سازد تا به سرعت و طی چند ساعت یا چند روز از یک ایده به کل مجموعه داده‌های یک موضوع برسند. محققان در حال حاضر می‌توانند توالی ژنگان ۵ نفر را در یک دور تعیین کنند و داده‌ها را در یک هفته و با هزینه‌ای کمتر برای هر ژنگان تولید کنند. اگر مقایسه کنیم، اولین ژنگان انسان با تکنولوژی CE حدود ۱۰ سال وقت نیاز داشت تا توالی‌یابی شود، بعلاوه به سه سال تجزیه و تحلیل نیاز داشت. مفهوم تکنولوژی NGS اساساً شبیه به CE است، بازهای قطعات کوچک DNA بوسیله سیگنال‌های منتشر شده در زمان سنتز هر قطعه از روی رشته DNA الگو تشخیص داده می‌شوند. NGS این فرایند را به جای یک یا تعداد کمی از قطعات DNA، با میلیون‌ها واکنش موازی انجام می‌دهد. این پیشرفت، تعیین توالی سریع مناطق بزرگی از جفت بازهای DNA در سراسر ژنگان را مقدور ساخته است. در پایان روش‌های متفاوت توالی‌یابی را نسبت به یکدیگر مقایسه می‌کنیم:

روش / ویژگی	Sanger	Solid	Illumina	454	Ion semiconductor	Single molecule real time sequencing
طول خوانش (bp)	۹۰۰–۴۰۰	۵۰–۳۵	۲۵۰–۵۰	۷۰۰	۲۰۰	۲۹۰
میزان صحت	٪۹۹,۹	٪۹۹,۹	٪۹۸	٪۹۹,۹	٪۹۸	٪۸۷
مقدار خوانش در هر دوره (Run)	نامحدود	۱,۲ تا ۱,۴ میلیون	بیش از ۳ میلیون	۱ میلیون	بیش از ۵ میلیون	۳۵–۷۵۰۰۰
زمان هر دوره	۳ ساعت و ۲۰ دقیقه	۱ تا ۲ هفته	۱ تا ۱۰ روز	۲۴ ساعت	۲ ساعت	۲ ساعت و ۳۰ دقیقه
مزایا	خواندن توالی های بلند، بسیار پرکاربرد	هزینه کمتر	پرقدرت، بازده بالا	سریع، خوانده توالی های بلند	سریع، هزینه کمتر	سریع، خواندن توالی های بلند
معایب	بسیار گران و غیر کاربردی برای پروژه های توالی یابی بزرگ	آهسته تر از روش های دیگر	مجهز و خیلی گران	خطاهای هموپلیمری، بسیار گران قیمت حتی در یک دوره	خطاهای هموپلیمری	گران، بازده کم

منابع

1. Hui, P. (2014). Next Generation Sequencing: Chemistry, Technology and Applications, Topics in Current Chemistry, Vol. 336, PP. 1-18.
2. Brown, T. A. (2010). Gene cloning and DNA analysis: an introduction, 6nd Edition, Wily –Black well press, New Jersey, United States, PP. 165- 84.
3. Heather, J. M., Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA, Genomics, Vol. 107, No. 1, PP. 1-8.
4. Gaastra, W. (1985). Chemical cleavage (maxam and gilbert) method for DNA sequence determination, Methods in molecular biology, Vol. 2, PP. 333-341.
5. Ansorge, W. J. (2009). Next-generation DNA sequencing techniques, New Biotechnology, Vol. 25, No. 4, PP. 195-203.
6. Huse, S. M., Huber, J. A., Morrison, H. G., Sogin, M. L., Welch, D. M. (2007). Accuracy and quality of massively parallel DNA pyrosequencing, Genome Biology, Vol. 8, No. 7, PP. 143.
7. Escalante, A. E., Barbolla, L. J., Ramírez-Barahona, S., Eguiarte, L. E. (2014). The study of biodiversity in the era of massive sequencing, Revista Mexicana de Biodiversidad, Vol. 85, No. 4, PP. 1249-1264.
8 Gupta, A. K., Gupta, U. D. (2014). Next Generation Sequencing and Its Applications, Animal Biotechnology, PP. 345-367.
9. Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics, Trends in genetics, Vol. 24, No. 3, PP. 133-141.
10. Anderson, M. W., Schrijver, I. (2010). Next Generation DNA Sequencing and the Future of Genomic Medicine, Genes, Vol. 1, PP. 38-69.
11. Diaz-Sanchez, S., Hanning, I., Pendleton, S., D'Souza, D. (2013). Next-generation sequencing: The future of molecular genetics, Poultry Science, Vol. 92, No. 2, PP. 562–572.
12. Mardis, E. R. (2008). Next-generation DNA sequencing methods, Annual review of genomics and human genetics, Vol. 9, PP. 387-402.
13. Gupta, P. K. (2008). Single-molecule DNA sequencing technologies for future genomics research, Trends in biotechnology, Vol. 26, No. 11, PP. 602-611.
14. Pareek, C. S., Smoczynski, R., Tretyn, A. (2011). Sequencing technologies and genome sequencing, Journal of applied genetics, Vol. 52, No. 4, PP. 413-435.
15. Kaiser, J. (2008). DNA sequencing. A Plan to Capture Human Diversity in 1000 Genomes, Science, Vol. 319, No. 5862, PP. 395.
16. Xuan, J., Yu, Y., Qing, T., Guo, L., Shi, L. (2013). Next-generation sequencing in the clinic: promises and challenges, Cancer Letters, Vol. 340, No. 2, PP. 284-295.
17. Ossowski, S., Schneeberger, K., Clark, R. M., Lanz, C., Warthmann, N., Weigel, D. (2008). Sequencing of natural strains of Arabidopsis thaliana with short reads, Genome research, Vol. 18, No. 12, PP. 2024-2033.
18. Pease, J., Sooknanan, R. (2012). A rapid, directional RNA-seq library preparation workflow for Illumina[reg] sequencing. Nature Methods, Vol. 9, No. 310. PP. 1-2.
19. Furey, T. S. (2012). ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions, Nature Reviews Genetics, Vol. 13, No. 12, PP. 840-852.
20. Forshew, T., Murtaza, M., Parkinson, C., Gale, D., Tsui, DW., Kaper, F., et al. (2012). Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA, Science translational medicine, Vol. 4, No. 136, PP. 136-168.