

گام های بزرگ در تعیین توالی DNA

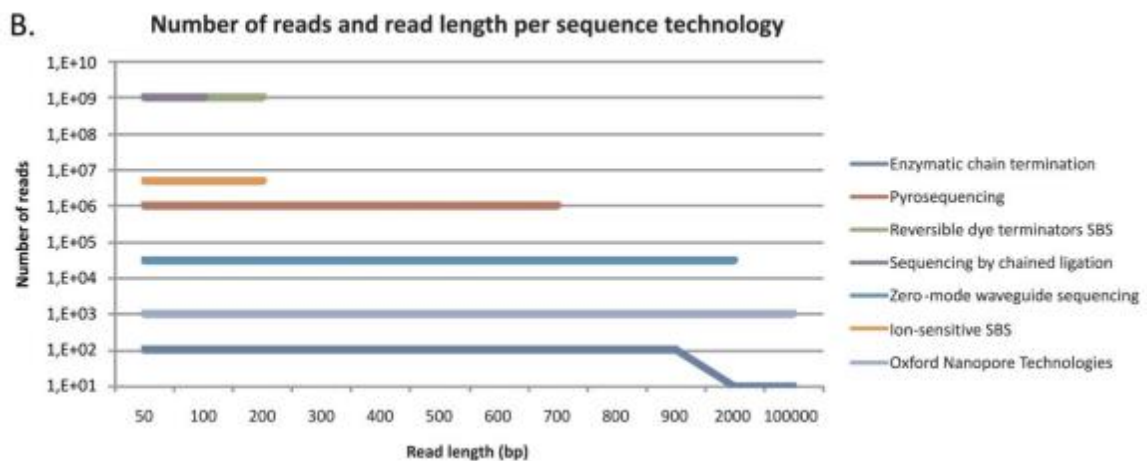
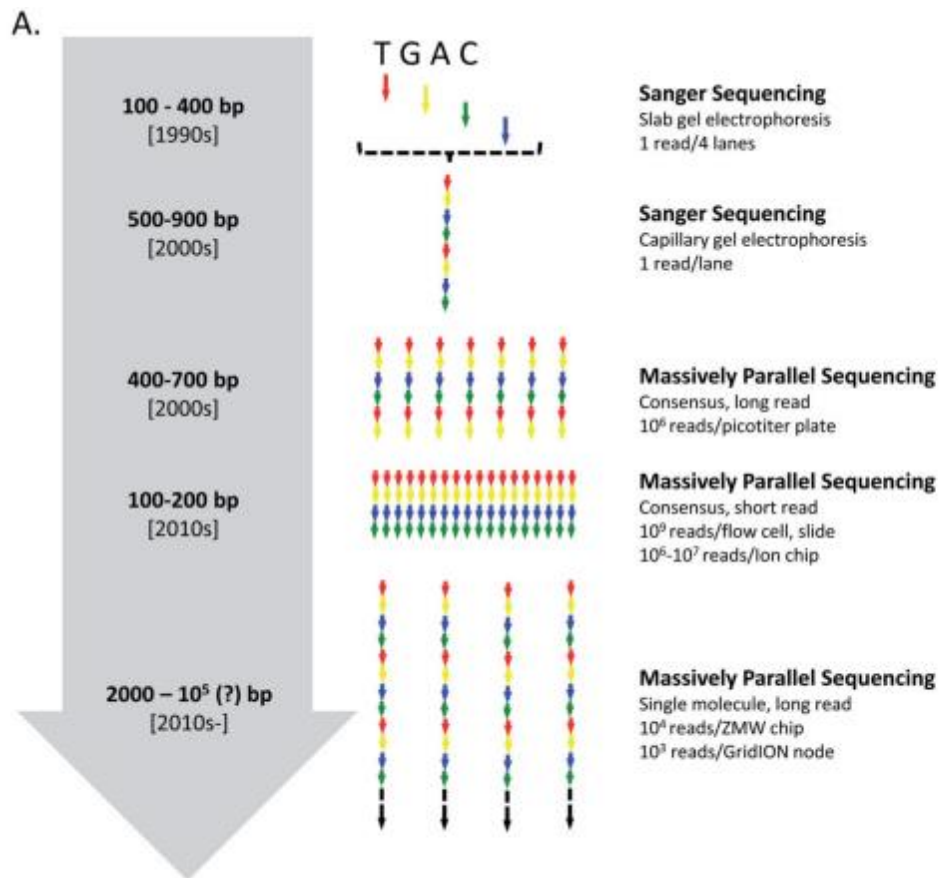
در سال های اخیر پیشرفت های فوق العاده ای در توانایی ما برای تعیین سریع و کم هزینه توالی DNA توالی به وجود آمده است. این مساله تحولات بنیادین در زمینه های ژنتیک و زیست شناسی ایجاد کرده که منجر به درک عمیق تر از رویدادهای مولکولی در فرایندهای زیستی شده است. پیشرفت های سریع تکنولوژیکی، فرصت ها و کاربردهای تعیین توالی را به سرعت گسترش داده است و در عین حال چالش هایی در مراحل پیشین تعیین توالی و در فرآیندهای پایین دستی بررسی و تجزیه و تحلیل این حجم انبوه از داده های توالی تحمیل کرده است. به طور سنتی، تعیین توالی، به قطعات کوچک DNA برای حدود یک هزار پایه (مشتق شده از ژنوم ارگانسیم) بمنظور حفظ کیفیت توالی و دقت بالا برای خواندن طول های بالاتر محدود شده است. اگر چه پیشرفت های تکنولوژیکی بسیاری بوجود آمده است، روش های تعیین توالی موازی که بطور تجاری در حال حاضر در دسترس هستند قادر به حل این مساله نیستند. روش تعیین توالی قادر به حل این مسئله نبوده است. با این حال، اعلان اخیر در تعیین توالی نانوحفره وعده از بین بردن این محدودیت خواندن طول را داده است، که امکان تعیین توالی قطعات سالم بزرگ تر DNA را می دهد. توانایی تعیین توالی DNA سالم با دقت بالا گامی بزرگ به سوی تسهیل تجزیه و تحلیل پایین دستی و افزایش قدرت تعیین توالی نسبت به امروز است. این بررسی برخی از پیشرفت های فنی در تعیین توالی که مرزهای جدیدی را در ژنتیک باز کرده اند پوشش می دهد.

کلمات کلیدی: DNA، ژنومیک، توالی نانوحفره، توالی موازی، توالی

1. مقدمه

درک کامل زبان DNA نیاز به تعیین کامل ترتیب پایه‌های ژنوم انسان (و یا دیگر موجودات مورد نظر) دارد. دستیابی به آن دانش، وعده بینش کامل‌تر نسبت به تغییرات بیولوژیکی و علل بیماری‌ها را می‌دهد. در سپیده دم تعیین توالی، خواندن ترتیب چهار پایه DNA یک فرایند دست و پا گیر بود. تعیین توالی اگرچه هنوز به سختی، با معرفی روش تجزیه شیمیایی صنعتی توسط Maxam و Gilbert [1] و روش تعیین توالی زنجیره پایانی که توسط sanger ساخته شده بود [2] در پایان دهه 1970 ممکن شد. ثابت شد روش دوم مفیدتر است و روش غالب تعیین توالی DNA برای تقریباً سه دهه بود و پروژه ژنوم انسان (HGP) را به پیش برد، و هنوز هم توسط بسیاری به عنوان "استاندارد طلایی" در نظر گرفته می‌شود. راه اندازی تجاری ابزارهای تعیین توالی موازی فشرده DNA در سال 2005 آغاز تغییر پارادایم طراحی شده توسط تکنیک‌های جدید تعیین توالی DNA بود که به محققان برای پاسخ به سوالات جسورانه در آزمایشات وسیع ژنوم الهام بخشید. در سال‌های اخیر رقبای جدید بسیاری در زمینه تعیین توالی موازی فشرده دیده می‌شود. قابل ذکر است، روش‌های تعیین توالی نوآورانه با استفاده از تک مولکول‌های DNA و تشخیص بی‌درنگ ظهور کرده‌اند. در حال حاضر، این رویکردهای نوآورانه در حال تکمیل پلتفرم‌های تعیین توالی موجود است، اما پیش از آن، آنها مسیری طولانی برای جایگزین شدن با روش تعیین توالی موازی فشرده دارند.

فن آوری‌های معمول تعیین توالی DNA و چالش‌های خود و محدودیت‌هایشان، در زیر توضیح داده می‌شود. خلاصه‌ای از ویژگی‌های هر یک از تکنولوژی‌ها در جدول 1 ارائه شده است حالی که تعدادی از خوانش توالی‌ها و خوانش طول در شکل 1 نشان داده شده است.



شکل 1. (A) تغییرات خواندن طول و درجه همسانی در فن‌آوری‌های تعیین توالی از دهه 1990 تا حالا. (B)

تعداد خوانش و طول هر خوانش برای تکنولوژی توالی

جدول 1. بررسی اجمالی تکنیک‌های تعیین توالی و مشخصاتشان

		Sequencing principle	Sample preparation	Detection principle	Major error modality	Read length (bp)	Sample throughput
First generation sequencing	Sanger sequencing	Enzymatic chain termination	Cloning/PCR	Optical	Substitutions	900	Low
Consensus sequencing	454	Pyrosequencing	Emulsion PCR	Optical	Indels	700	Medium
	Illumina/Solexa	Reversible dye terminators SBS	Bridge amplification	Optical	Substitutions	100	High
		SOLiD	Sequencing by ligation	Emulsion PCR	Optical	Substitutions	75
	Complete Genomics	Sequencing by ligation	Restriction/circularization/RCA	Optical	NA	35	High
	Ion Torrent	Ion-sensitive SBS	Emulsion PCR	Electronic	Indels	200	Medium
Single-molecule sequencing	Helicos	Reversible single-dye terminators SBS	No amplification	Optical	Deletions	35	Medium
	Pacific Biosciences	ZMW sequencing	No amplification	Optical	Indels	>2000	Medium
	Nanopore sequencing	Ionic current shift	NA	Electronic/Optical	NA	NA	NA

a) NA, not available; RCA, rolling circle amplification.

2. توالی sanger – توالی خاتمه دهی زنجیره

مهارت روش خاتمه‌دهی زنجیره در استفاده از نوکلئوتیدهای خاتمه زنجیره، دیدئوکسی- نوکلئوتید است که فاقد یک گروه 3' پریم هیدروکسیل است و بسط و گسترش بیشتر زنجیره DNA کپی شده را محدود می‌کند. برای تعیین توالی اولیه sanger نیاز است فرآیند توالی به چهار واکنش جداگانه تقسیم شود. هر واکنش شامل یک الگوی تک رشته‌ای DNA، یک آغازگر DNA و یک پلیمراز DNA در حضور مخلوطی از چهار نوکلئوتید اصلاح نشده، که یکی از آنها نشانه گذاری شده، و یک نوع نوکلئوتید خاتمه زنجیره اصلاح شده است. قطعاتی با طول متفاوت پس از هیبریداسیون پرایمر و اکتشن پلیمراز سنتز می‌شوند و همه همان 5' پریمپایانی را دارند اما توسط یک نوکلئوتید خاتمه زنجیره به 3' ختم می‌شوند. اضافه کردن تنها بخشی از نوکلئوتید پایان‌دار ترکیب تصادفی دیدئوکسی نوکلئوتیدها را تنها در بخش کوچکی از مولکول‌ها تضمین می‌کند. از آنجا که نوکلئوتیدهای خاتمه زنجیره مختلف در چهار واکنش توالی یابی استفاده می‌شود، تمام ترکیبات خاتمه می‌تواند تولید شود. الگوهای DNA محدود به 3' تولید شده پس از تغییر ماهیت بوسیله حرارت و تجزیه و تفکیک شدن توسط ژل الکتروفورز، فرآورده‌های هر چهار واکنش تعیین توالی موازی را ادامه می‌دهد [2]. با استفاده از رادیواکتیو یا،

اخیراً، بوسیله برچسب زنی با فلورسنت برای تجسم باندها دنباله‌ای از الگوهای DNA اصلی را قادر می‌سازد تا با پیروی از ترتیب مهاجرت قطعات بزرگتر متوالی در ژل تعیین شود.

چند پیشرفت در روش اصلی توسعه یافته توسط sanger بوجود آمده است. اینها عبارتند از: برچسب زنی خاتمه زنجیره نوکلئوتیدها با رنگهای فلورسنت طیفی مجزا که قادر می‌سازد: هر لوله و خط تنها در گام‌های تولید و شکستن قطعه مورد استفاده قرار گیرد [3، 4]؛ حذف نیاز به کست ژل با استفاده از الکتروفورز ژل مویرگی [5، 6]؛ و اتوماسیون پروتکل، که منجر به افزایش موازی‌سازی، تکرارپذیری و توان عملیاتی می‌شود [7].

توالی sanger هنوز به طور گسترده‌ای امروز برای بسیاری از کاربردها، به خصوص اعتبارسنجی متغیرهای ژنتیکی و در مواردی که کیفیت بالای خوانش 300-900 پایه مورد نیاز است استفاده می‌شود. با این حال، پیشرفت‌های عمده‌ای در تکنولوژی توالی در سال‌های اخیر به روش توالی sanger منتسب نشده است، اما روش‌های تعیین توالی موازی فشرده به سرعت در حال پدیدار شدن است.

3. توالی موازی فشرده - توالی اجماع

پلتفرم‌های تعیین توالی که امروزه به طور گسترده‌ای در تحقیقات ژنتیکی مورد استفاده قرار می‌گیرند، پلتفرم‌های تعیین توالی فشرده هستند. که ویژگی‌های بسیاری از توالی sanger به ارث برده‌اند مانند استفاده از پلیمرها برای سنتز، نوکلئوتیدهای اصلاح شده و تشخیص فلورسنت. ویژگی دیگر این است که آنها نیاز به DNA دارند تا بطور کلونال تقویت شوند تا تشکیل یک الگوی اجماعی قبل از تعیین توالی را بدهند. این روش‌های تعیین توالی به نام نسل بعدی تعیین توالی خوانده می‌شوند، روش‌های تعیین توالی با توان عملیاتی بالا، و یا نسل دوم تعیین توالی. با این حال، همچنانکه تکنولوژی‌های تعیین توالی توسعه می‌یابند، حفظ رجوع به فناوری‌های تعیین توالی قدیمی و بدیع در حال ظهور متعلق به یک نسل خاص سخت است. بنابراین، بهتر است آنها را بر حسب برجسته‌ترین ویژگی‌های مشترکشان، مانند روش‌های تعیین توالی موازی فشرده و یا تک مولکولی گروه‌بندی کنیم. در حال حاضر، پنج رقیب در زمینه فناوری‌های تعیین توالی موازی فشرده وجود دارد، هر کدام با نقاط قوت و ضعف خاص. این‌ها در ادامه توصیف و مورد بحث قرار گرفته‌اند.

3.1. تعیین توالی به روش پیرو

Melamede در ابتدا مفهوم تعیین توالی با سنتز (SBS) را در سال 1985 در یک گزارش از تلاش‌ها برای تشخیص رویداد ترکیب نوکلئوتید با اندازه‌گیری جذب نوکلئوتید [8] را مطرح کرد. بی‌خبر از یافته‌های Nyrén ، Melamede رویکرد SBS دیگری را با استفاده از زیست‌تابی به جای جذب در سال 1986 تصور کرد [9]، که منجر (12 سال پس از آزمایش دیگری) به معرفی pyro-sequencing شد [10]. فن‌آوری pyrosequencing قادر به تکمیل توالی sanger بود و بیشتر توسط علوم زندگی 454، که توسط جاناتان روتبرگ پایه‌گذاری شد، توسعه داده شد. در سال 2005، روتبرگ و همکارانش [11] اولین مقاله اثبات مفهومی را منتشر کردند که نشان داد رویکرد pyrosequencing در تعیین توالی موازی فشرده، بازده فوق‌العاده‌ای در افزایش ظرفیت توالی، و در نتیجه دگرگونی روش تعیین توالی دارد.

در SBS، این رویداد است که تشخیص داده می‌شود ترکیب نوکلئوتید با رشته‌های DNA در حال رشد است. همچنانکه نوکلئوتیدها توسط پلیمراز ترکیب می‌شوند، پیرو فسفات (PPI) و پروتون‌ها تولید می‌شوند. در pyrosequencing، PPI در یک آبشار آنزیمی برای تولید یک انفجار نوری مورد استفاده قرار می‌گیرد. PPI بوسیله سولفوریلاز ATP به ATP تبدیل می‌شود، که سپس از آن برای تولید فوتون توسط لوسیفراز آنزیم کرم شب تاب آنزیم استفاده می‌شود، که یک سیگنال نوری متناسب با تعداد نوکلئوتیدهایی که گنجاینده شده است فراهم می‌کند. چالش، تشخیص چشمک‌های نور از هر الگوی DNA منحصر به فرد است هنگام تعیین توالی الگوهای متعدد به صورت موازی است. انفصال فضایی هر واکنش توالی بر روی مهره‌های ته نشین شده در ته چاه کوچک بر روی یک صفحه picotiter به زیبایی این مشکل را حل کرد.

pyrosequencing موازی فشرده با قطعه قطعه شدن DNA و اتصال آداپتوری آغاز می‌شود. الگوهای تک رشته‌ای DNA سپس بر روی دانه‌ها محدود می‌شوند و امولسیون PCR انجام می‌شود، بطور کلونالی هر الگوی DNA در میکرواکتورهای آبی جدا شده با روغن تقویت می‌شود. امولسیون سپس شکسته می‌شود و دانه‌های حامل DNA از دانه‌های خالی در فرآیندی به نام غنی سازی از هم جدا می‌شوند. دانه‌های غنی در چاه‌های کوچک در صفحه picotiter با پرایمر و پلیمراز DNA پلیمرز با هم ته‌نشین می‌شوند. در حالت ایده آل، تنها یکی از این دانه‌ها برای هر چاه مناسب خواهد بود. دانه‌های کوچکتر نیز اضافه می‌شوند، که آنزیم‌های مسئول

تولید نور با استفاده از PPI را حمل می‌کنند. نوکلئوتیدها پس از آن از روی زیر لایه در یک جریان آرام از محلول مورد استفاده در یک ترتیب از پیش تعیین شده عبور کرد، و چشمک‌های نور در هر چاه متناظر با ترکیب ثبت شد. حذف کارآمد واکنش با محصولات، که در صورت عدم حذف می‌تواند واکنش توالی را آشفته کند، توسط جریان آرام تسهیل می‌شود. از این رو، الگوهای توالی DNA با آگاهی از محل آنها تعیین می‌شود، منظور جریان نوکلئوتید و سوابق هر یک از فلش‌های نور از هر چاه است [10-12]. اشکال عمده استفاده از این رویکرد، مشکلات در بسط توالی نوکلئوتیدهای یکسان (مناطق homopolymeric) طولانی‌تر از حدود پنج نوکلئوتید با توجه به پاسخ غیرخطی نوری است که آن‌ها تولید می‌کنند [10].

طول خوانش قابل مقایسه و افزایش ظرفیت رشته در مقایسه با توالی سنتی sanger تعیین توالی موازی فشرده با استفاده از pyrosequencing را برای توالی de novo، تعیین دوباره توالی ژنوم و مطالعات متاژنومیک ایده‌آل کرده است [12]. جیمز واتسون، که به حل ساختار DNA کمک کرده است، ژنومش را در سال 2008 با استفاده از pyrosequencing تعیین توالی کرده بود [13]. مهمترین اثر این پلتفرم توالی را می‌توان با مقایسه توالی ژنوم واتسون - تنها در دو ماه با هزینه 1 میلیون دلاری تکمیل شد [13] - با تلاش‌های HGP، در زمان 11 سال و هزینه 3 میلیارد دلاری متوجه شد [14]. نقطه عطف دیگر با استفاده از pyrosequencing در تعیین توالی موازی فشرده کاری است که در ژنوم ناناندرتال توسط پابو و همکاران انجام شد [15]. به تازگی، طول خوانش پلتفرم از 450 پایه به 700 پایه توالی معمول sanger افزایش یافته است. تعیین توالی موازی فشرده با pyrosequencing بعلاوه افزایش توان عملیاتی بر توالی sanger بسیار برتری یافته است. با این حال، جامعه پژوهش به نظر می‌رسید اشتباهی سیری ناپذیری را برای داده‌های توالی است توسعه داده است که به راحتی در مورد استفاده از زیست‌تابش SBS راضی نمی‌شدند.

3.2 SBS نابودگر رنگ برگشت پذیر

سود اساسی در توان عملیاتی تعیین توالی زمانی که دو سیستم تعیین توالی موازی فشرده برای مصارف تجاری در سال 2006 راه اندازی شد به دست آمد. این تکنولوژی مبتنی بر SBS با استفاده از رنگ پایان دهی

نوکلئوتید برگشت‌پذیر چهار طیفی است و برای اولین بار در سال 2001 توسط یک شرکت کوچک، Solexa، [16، 17]، که بعدها توسط ILLUMINA [18] به دست آمد تعیین شده بود.

SBS با استفاده از پایان دهی رنگ قابل برگشت با تکه تکه شدن DNA و اتصال آداپتورها آغاز می‌شود. به جای استفاده از امولسیون PCR برای تفکیک فضایی هر الگوی DNA، در رویکرد pyrosequencing تعیین توالی موازی فشرده، روش دیگری بکار گرفته می‌شود. الگوهای DNA از طریق آداپتورها به یک سطح مسطح هیبریده می‌شوند، جایی که در آن هر الگوی DNA بطور کلونال توسط PCR فاز جامد تقویت می‌شود، همچنین به عنوان تقویت پل شناخته می‌شود. این، سطحی را با یک چگالی بالا از خوشه‌های فضایی مجزا ایجاد می‌کند، که هر خوشه شامل یک الگوی DNA منحصر به فرد است. این‌ها در ابتدا با عبور پایان دهنده‌های رنگ برگشت‌پذیر چهار طیفی مجزا در یک جریان محلول بر روی سطح در حضور یک پلیمراز DNA تعیین توالی می‌شوند. فقط گسترش‌های تک پایه‌ای با توجه به اصلاح پریم 3 از نوکلئوتیدهای پایان دهی زنجیره ممکن است، و هر خوشه شامل فقط یک نوع از نوکلئوتیدهاست، همچنانکه بوسیله الگوی DNA تشکیل دهنده خوشه دیکته شده است. پایه گنجانده شده در تمام خوشه‌ها توسط تصویربرداری فلورسانس از سطح قبل از حذف شیمیایی رنگ و ترمیناتور شناسایی می‌شود، پایه قابل گسترش که برای دور جدیدی از توالی آماده است تولید می‌شود. رایج ترین خطاهای توالی تولید شده در SBS پایان دهی رنگ برگشت پذیر جانشینی‌ها هستند [18، 19].

یکی از اشکالات عمده شیمی پایان دهی رنگ برگشت پذیر محدودیت در طول خواندن است، بطوریکه رسیدن به بهره وری 100٪ از اختلاط پایه‌ها و شکافتگی در هر چرخه دشوار است. هنگامی که سیستم راه اندازی شد، کیفیت بالای خواندن طول تنها 30-36 پایه را نمایان کرد [18، 20]، که به طور چشمگیری کمتر از طول خواندن ارائه شده توسط هر دو توالی sanger و pyrosequencing برای تعیین توالی موازی فشرده بود. در واقع، آن حتی کمتر از طول خواندن روش "plus-minus" که در سال 1975 بوسیله سنگر و کولسون توسعه داده شد بود [21]، اما یک افزایش چشمگیر در خروجی داده‌های توالی این اشکال را جبران کرد. پیشرفت‌های قابل توجهی در هر دو زمینه طول خواندن و توان عملیاتی از زمان معرفی به دست آمده است، که SBS پایان دهی رنگ برگشت پذیر را قادر می‌سازد تبدیل به موفقترین پلتفرم تعیین توالی موازی فشرده شود. طول

خواندن کوتاه در ابتدا استفاده از آن در اپلیکیشن‌های اسمبلی بزرگ *de novo* را ممنوع کرد، اما آن را برای توالی مجدد و اپلیکیشن‌های شمارش داده، به عنوان مثال توالی ترنسکریپتومیس، *RNA-SEQ* و تعیین توالی قطعات *DNA* که در فعل و انفعالات پروتئین *DNA* درگیر هستند، *ChIP-Seq* مناسب کرده است [18]. بهبود در طول خواندن و توسعه الگوریتم‌های اسمبلی *de novo*، مانند *SOAP de novo* [22] در حال حاضر استفاده از *SBS* پایان دهی رنگ برگشت پذیر را در اسمبلی *de novo* ژنوم پستانداران بزرگ، به عنوان مثال، ژنوم پاندا را امکان پذیر کرده است [23]. به تازگی، *ILLUMINA* سیستم سریع *2500Hiseq* خود را ارائه کرده است که قادر به تعیین توالی ژنوم انسان در 27 ساعت، و فعال کردن طول خواندن اطلاعات تا 150 پایه از طریق چرخه‌های زمانی کوتاه‌تر توالی است.

3.3. توالی بوسیله زنجیره لیگاسیون

در طبیعت، لیگازهای *DNA* آنزیم‌های ضروری برای تعمیر شکاف‌ها و شکستگی‌های *DNA* هستند و در واکنش‌های ترمیم و همانندسازی *DNA* نقش دارند [24]. در سال 2005، *Church and Shendure* نشان دادند که واکنشی که آن‌ها کاتالیز می‌کنند می‌تواند برای توالی قطعات *DNA* کوتاه ساکن بر روی دانه‌های به دام افتاده در ژل پلی آکریل امید در فرآیندی به نام توالی *polony* استفاده شود [25]. *SOLiD* (*Sequencing by Oligonucleotide Ligation and Detection*) سومین پلتفرم تجاری توالی موازی فشرده، بر اساس یک نسخه پیچیده تر از این روش، در سال 2007 به بازار 2007 آمده است [26].

در ابتدا، پلتفرم *SOLiD* بسیاری از پایه‌های پلتفرم *ILLUMINA* را به همراه داشت، اما ارتقاءهای اخیر در مقابل *2000Hiseq chemistry*، را عملکرد این ابزار را به طور قابل توجهی بالاتر از *SOLiD* برده است با این حال، در مراحل آماده سازی کتابخانه، پلتفرم *SOLiD* شبیه تعیین توالی فشرده موازی با *pyrosequencing* است. تقویت کلونال با استفاده از امولسیون *PCR* پس از ایجاد یک کتابخانه *adaptor-ligated* به دست می‌آید. دانه‌های غنی شده سپس روی اسلاید شیشه ای بطور تصادفی ته نشین شده با یک اصلاح 3 پریم به اهداف معرفی شده که پوشش دهنده دانه‌ها هستند متصل شدند. در توالی‌های بعدی با *chained* لیگاسیون،

الیگومرهای هشت نوکلئوتید، اکتامرها، به عنوان پروب تشخیص به جای تک نوکلئوتید استفاده می‌شود. هر اکتامر شامل دو پایه شناخته شده در انتهای 3 پریم است که با سه نوکلئوتید دژنره و سه نوکلئوتید دیگر که می‌تواند با هر هر نوکلئوتید دیگری هیبرید شود، به نام نوکلئوتید یونیورسال شناخته می‌شود. چهار رنگ بطور طیفی مجزا به کار گرفته می‌شوند، که هر کدام توسط چهار اکتامر حمل می‌شوند و یک مجموعه پروب متشکل از 16 اکتامر، تمام ترکیبات ممکن دو پایه شناخته شده را پوشش می‌دهند. رنگ به انتهای 5 پریم هر اکتامر متصل می‌شود. برای شروع چرخه توالی اول، دانه‌های پوشش دهنده DNA با یک پرایمر توالی ، 16 پروب اکتامر و لیگاز انکوبات می‌شوند. فقط یک پروب کاملاً هیبرید شده به پرایمر توالی توسط لیگاز خواهد پیوست. پروب ها دو پایه اول را رمزگشایی می‌کنند ، و سپس بطور شیمیایی شکافته خواهند شد تا آخرین سه پایه 5 پریم حذف شوند. این چرخه از هیبریداسیون و لیگاسیون پروب، تصویربرداری فلورسنت و شکافت شیمیایی ده بار تکرار می‌شود برای تحقیق در مورد تمام پایه‌های پنجم. قطعات ایجاد شده سپس از دانه‌ها تهی می‌شوند و یک دور لیگاسیون دوم با استفاده از توالی پرایمر انیلینگ یک پایه بالادست از قبلی انجام می‌شود. از این رو، هر پایه به دو رنگ فراخوانی مرتبط است. به تدریج تغییر موقعیت پرایمر توالی و تکرار بستن دور لیگاسیون با استفاده از ده چرخه در هر بار منجر به فراخوانی رنگ می‌شود، که می‌تواند برای به دست آوردن یک دنباله خطی در فضای رنگ رمزگشایی شود. ترجمه از فضای رنگ به یک توالی نوکلئوتیدی عادی با بیشترین دقت بوسیله همترازی خوانش فضای رنگ به یک ژنوم مرجع فضای رنگ انجام می‌شود. به عنوان یک نتیجه از همترازی فضای رنگ، اشتباهات می‌تواند بطور موثری اصلاح شود از آنجایی که تنها ترکیبات رنگی خاصی مجاز است، تسهیل تمایز بین جانشینی نوکلئوتید، نمایانگر یک تغییر دو رنگه در نوکلئوتیدهای مجاور و اشتباهات توالی است [28-26]. رایج‌ترین خطاها در تعیین توالی با زنجیره لیگاسیون قابل تعویض هستند [18].

Life Technologies به تازگی یک جایگزین برای مرحله پرزحمت آماده سازی کتابخانه امولسیون با یک تکنیک به نام wildfire ارائه کرده است که شبیه تقویت پل در پلت فرم ILLUMINA است که وعده کاهش زمان آماده سازی کتابخانه زمان آماده سازی و همچنین چگالی بالاتر روی سطح تراشه را می‌دهد که توان عملیاتی بالاتر را ممکن می‌سازد. سیستم SOLiD بهترین مورد برای پروژه‌های تعیین توالی است که نیاز به نرخ خطای پایین [26]، تعیین توالی ترانسکریپتوم و شمارش برچسب‌ها دارد، به عنوان مثال ChIP-Seq با

توجه به طول خواندن کوتاهش، سیستم تصحیح خطا و خروجی انبوهی از داده‌ها [18]. رویکرد فضای رنگی بطور موثری خطاهای توالی را آشکار می‌کند، اما ثابت شده است تجزیه و تحلیل داده‌های پایین دستی دست و پا گیر است که منجر به توسعه چند ابزار اپن سورس مرتبط شده است [29]. با این حال، بهره برداری از واکنش لیگاسیون در توالی بدون استفاده از رویکرد فضای رنگی ممکن است.

3.4. توالی با استفاده از لیگاسیون حذف زنجیره

در سال 2010، Complete Genomics Inc یک مقاله چاپ کرد و قابلیت‌هایش در تعیین توالی کامل ژنوم انسان با استفاده از متد تعیین توالی با استفاده از روش به اشتراک گذاری چندین ویژگی با توالی polony را نشان داد [25، 30]. به جای فروش ابزارهای تعیین توالی، که تاکنون مدل کسب و کار غالب بوده است، شرکت خدمات برون سپاری برای تعیین توالی DNA انسان شامل آماده سازی کتابخانه، تعیین توالی و تجزیه و تحلیل عمومی را فراهم می‌کند.

Complete Genomics تعیین توالی لیگاسیون حذف زنجیره را با استفاده از فن‌آوری ترکیبی خود probe-anchor ligation (cPAL) انجام می‌دهد. DNA برای تعیین توالی به 500 پایه خرد می‌شود و چهار آداپتور به هر بخش از ژنومیک از طریق برش‌های مکرر با آنزیم‌های محدود و لیگاسیون درون مولکولی فرستاده می‌شود. چرخه‌های منتهی پس از آن در محلول به رول‌های تک رشته‌ای DNA تقویت می‌شوند [31]، که نانوتوپ‌های DNA نامیده می‌شود، که متشکل از نسخه‌های مکرر چرخه‌های اصلی است. نانوتوپ‌های DNA سپس به طور تصادفی به نقاط دستور داده شده بر روی یک سطح مسطح متصل می‌شوند. هر نقطه به اندازه کافی کوچک فقط شامل یک نانوتوپ DNA است. رویکرد Complete Genomics در تعیین توالی با لیگاسیون حذف زنجیره از یک پروب لنگری برای پیدا کردن موقعیت آداپتورها و پروب تشخیص شامل نه پایه (nonamers) و چهار رنگ استفاده می‌کند. پروب لنگری و پروب تشخیص توسط هیبریداسیون و لیگاسیون متصل شده‌اند تا یک پایه هر چرخه را رمزگشایی کنند. نانوتوپ‌ها می‌توانند رمزگشایی شوند، یک نوکلئوتید در هر لحظه، در هر دو طرف موقعیت آداپتور. به طور خلاصه، یک الیگونوکلئوتید لنگری با یکی از چهار آداپتور هیبرید می‌شود، و چهار پروب، هر کدام حامل یک رنگ طیفی مجزا حمل می‌کنند برای تحقیق از یک موقعیت

پایه مجاور به آداپتور استفاده می‌شود. رنگ به یک پایه شناخته شده در یک وضعیت خاص در پروب نونامر متصل می‌شود در حالی که بقیه موقعیت‌ها توسط نوکلئوتیدهای دژنره پر شده است. تغییر موقعیت پایه شناخته شده در پروب موقعیت‌های 1-5 در الگوی DNA اجازه می‌دهد تا در مجاورت آداپتور خوانده شوند. ایجاد یک پروب لنگری بسط یافته با لیگاسیون دو پروب لنگری اجازه می‌دهد تا رمزگشایی از مواضع 6-10 مجاور آداپتور انجام شود. مجموعه لنگر- پروب پس از اشکارسازی فلورسنت هر هیبریداسیون و واکنش لیگاسیون تکمیل می‌شود. این سیستم را بازنشانی می‌کند. از این رو، هر فراخوانی پایه از زنجیره رها می‌شود، در نتیجه کیفیت توالی بهبود می‌یابد زیرا هر پایه رمزگشایی شده مستقل از کامل بودن چرخه‌های قبلی است. کل پروسه سپس برای هر دو طرف هر سایت آداپتور تکرار می‌شود، و در نتیجه 10-mers ها در خوانش 35 پایه-زوج- پایان ادغام می‌شوند [30].

دلایلی که Complete Genomics در حال نگه داشتن پلتفرم توالی آنها در خانه هستند، حداقل بطور جزئی، مربوط به چالش قابل توجه نقشه برداری و مونتاژ ژنوم انسان بسیار تکراری با استفاده از خوانش 35-پایه-زوج-پایان کوتاه و نسبتاً خواستار آماده سازی کتابخانه cPAL است. ILLUMINA به ابتکار Complete Genomics با تعیین توالی کامل ژنوم انسان به عنوان یک سرویس با استفاده از شیمی پایان دهی رنگ بازگشت پذیر پاسخ داده است. آن باقی می‌ماند تا دیده شود آیا بیشترین تعیین توالی در بیمارستان‌ها و مراکز تعیین توالی انجام خواهد شد، و یا به شرکت‌های تعیین توالی برون سپاری خواهد شد.

3.5 حساس یونی SBS

تمام پلتفرم‌های تعیین توالی موازی فشرده تاکنون از اشکارساز نوری برای کشف یک توالی DNA استفاده می‌کنند. تبادل تشخیص نوری با تشخیص الکتریکی به طور بالقوه چندین مزیت دارد چرا که این امر از نیاز به نوکلئوتید اصلاح شده، اپتیک گران و اکتساب وقت گیر و پردازش مقادیر زیادی از داده‌های تصویر فلورسنت جلوگیری می‌کند. از این رو، آن وعده تعیین توالی ارزان تر، کارآمدی فضایی و سرعت بیشتر را می‌دهد. پورمند و همکارانش [32] یک روش SBS تشخیص الکتریکی در سال 2006 ارائه کردند، که نشان می‌دهد که رویدادهای سنتز DNA را می‌توان با استفاده از آشفتگی بار الکتریکی با استفاده از یک تقویت کننده ولتاژ

جهشی، شناسایی کرد. در سال 2010، روتبرگ و همکارانش [33] یک پلتفرم ابتکاری تعیین توالی بر اساس کار پورمند و همکارانش [34]، از جریان یونی که متکی بر سیگنال‌های الکتریکی به جای فلورسنت برای بدست آوردن اطلاعات توالی استفاده کردند.

این فرایند بسیار شبیه به آن چیزی است که قبل از آن توسط روتبرگ و همکاران برای تعیین توالی موازی فشرده با pyrosequencing توسعه یافته است. این شامل استفاده از همان مراحل آماده‌سازی کتابخانه، با یک کتابخانه قطعه adaptor-ligated است، که بطور کلونال بر روی دانه‌ها با استفاده از روش PCR امولسیون تقویت می‌شود. یک پلیمراز DNA برای ترکیب نوکلئوتیدها روی یک الگوی DNA پرایم شده با یک جریان از پیش تعیین شده از نوکلئوتیدها استفاده می‌شود. با این حال، تمرکز از PPI به دیگر محصول ترکیب نوکلئوتید، یون هیدروژن تغییر می‌یابد. دانه‌های غنی شده از امولسیون PCR در چاه‌های کوچک روی تراشه یونی ته نشین می‌شوند (هر یک مجهز به یک سنسور است که قادر به تشخیص پروتون آزاد است). همچنانکه پلیمراز DNA نوکلئوتید را به الگوی DNA اضافه می‌کند، یونهای هیدروژن آزاد می‌شوند، و در نتیجه تغییر مکان پروتون آزاد است توسط سنسور تشخیص داده شده و به یک سیگنال الکتریکی تبدیل می‌شود. تراشه یونی که مبتنی بر فناوری است که به طور گسترده برای ساخت مدارهای مجتمع استفاده می‌شود، تولید کم هزینه، و در مقیاس بزرگ را تسهیل می‌کند و باید مقیاس پذیری عالی را اعطا کند. با این حال، به مانند تعیین توالی موازی فشرده با pyrosequencing، نواحی homopolymeric منجر به افزایش حذف و الحاق، با توجه به پاسخ الکتریکی غیرخطی بر ترکیب بیش از پنج نوکلئوتید می‌شود [33].

تعیین توالی ژنوم گوردون مور [33] نشان داد که ممکن است، اما شاید استفاده از سیستم جریان یونی برای تعیین مجدد توالی پستانداران بزرگ قابل توصیه نباشد. سیستم جریان یونی دارای یک زمان اجرای سریع 2 ساعته است اما هنگامی که راه اندازی شد از نظر توان عملیاتی و طول خوانده شده شبیه pyrosequencing برای تعیین توالی موازی فشرده بود. ان نظر می‌رسد برای تعیین مجدد توالی de novo و تعیین توالی ژنوم باکتریایی و یا نواحی مورد هدف مناسب ژنوم موجودات پیچیده‌تر مناسب باشد. شیوع اخیر تولید شیگا توکسین از باکتری اشرشیاکلی یک نمونه عالی از ارزش یک پلتفرم تعیین توالی با خروجی نسبتاً کم ولی با زمان اجرای سریع توالی هنگام نظارت بر شیوع پاتوژن [35، 36] مهیا کرد. جانشین جریان یونی به نام پروتون یون شناخته

می‌شود، که برنامه ریزی شده تا در سال 2012 منتشر شود، و مبتنی بر همان شیمی سلف خود است اما انتظار می‌رود که دارای حدود 60 برابر چاه بیشتر روی تراشه پروتون یونش باشد. این افزایش عظیم در توان باید تعیین توالی ژنوم یک انسان در عرض چند ساعت از توالی را ممکن کند.

3.6. موضوعات جاری در تعیین توالی موازی فشرده

روش‌های تعیین توالی موازی فشرده از نظر بیوشیمی توالی کاملاً متنوع هستند، اما آن‌ها بسیاری از ویژگی‌های مشترک را به اشتراک می‌گذارند. مراحل آماده سازی کتابخانه‌شان با تکه تکه شدن تصادفی DNA که با لیگاسیون آداپتورهای پلتفرم خاص در پایان هر قطعه آغاز می‌شود. این آداپتورها سپس در تکثیر قطعه بر روی یک سطح جامد توسط پلیمرز استفاده می‌شود، به جز در تقویت cPAL است، که در محلول انجام می‌شود. محصولات تقویت بصورت فضایی در یک آرایه قبل از تعیین توالی خوشه بندی می‌شوند. روند تعیین توالی خودش با سری‌های خودکار هماهنگ بیوشیمیایی کنترل شده با انزیم و مراحل اکتساب داده با تصویربرداری فلورسنت انجام می‌شود. تنها جدیدترین سیستم، جریان یونی، بر اساس تغییرات مکانی پروتون آزاد، قادر به اشکارسازی الکتریکی است. همه این پلتفرم‌ها همچنین قابلیت خواندن هر دو به پایان یک قطعه DNA را دارند که تعیین توالی paired-end نامیده می‌شود. این ویژگی در حل نواحی تکراری در ژنوم و تعیین کیفیت ترانسکریپت ایزوفرم در RNA-SEQ سودمند است.

گام محدود کننده نرخ در روند تعیین توالی به طور سنتی واکنش تعیین توالی بوده است. با این حال، تغییر نسبت به پایان HGP آغاز شده است در حالیکه ظرفیت ابزارهای تعیین توالی شروع به تجاوز از نرخ کرده است که در آن نمونه‌های جدید برای تعیین توالی می‌تواند آماده شود. ابزارهای فعلی تعیین توالی چندین برابر بیشتر داده نسبت به توالی سانگر معمولی تولید می‌کنند، که مراحل محدود کننده سرعت را به آماده سازی کتابخانه و تجزیه و تحلیل داده‌ها شیف‌ت می‌دهد. چالش، حفظ روند آماده سازی نمونه، واکنش تعیین توالی و تجزیه و تحلیل داده‌های متعادل شده است. از این رو، اتوماسیون آماده سازی کتابخانه و مراحل کنترل کیفیت نقش حیاتی در حفظ سرعت با افزایش قدرت عملیاتی تعیین توالی ایفا می‌کند. طول‌های خواندن عموماً کوتاه همراه مقدار زیادی از داده‌هایی که باید تجزیه و تحلیل شوند و دقت خام کاهش یافته، در مقایسه با توالی سانگر،

چالش‌های بسیاری در تجزیه و تحلیل داده‌های پایین دستی [18] معرفی کرده است. با این حال، توسعه الگوریتم‌های جدید متناسب با انواع جدیدی از داده‌های تولید شده و استفاده از خوشه‌های ابر رایانه برای تجزیه و تحلیل داده‌ها برخی از این چالش‌ها را کاهش داده است.

4. تعیین توالی موازی فشرده - تعیین توالی تک مولکولی

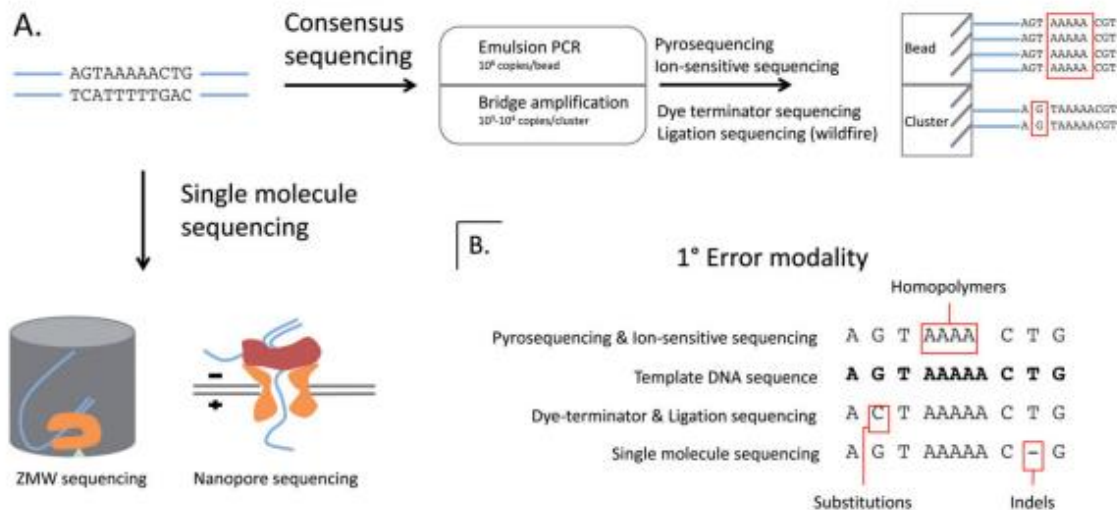
تعیین توالی اجماع موازی فشرده به تکنولوژی غالب در تعیین توالی تبدیل شده است، اما روش‌های دیگر که از تقویت الگوهای پیشین DNA برای تعیین توالی پرهیز می‌کنند ظهور کرده‌اند. هدف از این فناوری‌ها تعیین توالی با تک مولکول DNA و ترجیحا در بصورت بلادرنگ است. مزایای بالقوه استفاده از تعیین توالی تک مولکولی عبارتند از: مقادیر حداقل DNA ورودی مورد نیاز؛ حذف بایاس تقویت؛ سنتز اسنکرون؛ زمان چرخش سریع؛ و ظرفیت بررسی ویژگی‌های تک مولکول‌های DNA است. یک مقایسه از تعیین توالی اجماعی و تعیین توالی تک مولکولی، و همچنین رایج‌ترین اشتباهات برای هر یک از تکنولوژی‌های تعیین توالی در شکل 2 نشان داده شده است.

4.1. ترمیناتور تک رنگ برگشت پذیر SBS

Helicos Biosciences اولین تعیین توالی تک مولکولی DNA را در سال 2008 به نام سیستم Helicos معرفی کرد [37]. این سیستم از نوکلئوتیدهای پایان دهی زنجیره برگشت پذیر لیبل گذاری شده در یک رویکرد SBS با DNA بیحرکت بر روی یک سطح مسطح استفاده می‌کند [38, 39]. از این رو، بسیاری از ویژگی‌ها را با توالی اجماعی موازی فشرده در برپایی، بیوشیمی، توالی زمان اجرا و توان به اشتراک می‌گذارد. آن در اولین توالی تک مولکولی ژنوم انسان در سال 2009 استفاده شد [40] و می‌تواند برای تقویت توالی RNA استفاده شود [41]. با این حال، پلتفرم از سال 2008 است در بازار بوده است و به نظر می‌رسد در حال مبارزه در رقابتی شدید است. دلایل احتمالا در نرخ خطای خواندن خام بالای آن ($< 5\%$) و طول کوتاه خواندن (~ 32 پایه) است [42].

4.2. تعیین توالی موجبر حالت صفر با پلیمرز تثبیت شده

Pacific Biosciences اولین پلتفرم توالی SBS تک مولکولی بلادرنگ را در یک مقاله در Science در سال 2009 ارائه کرد. پلتفرم شامل چندین ویژگی ابتکاری جالب، از جمله تثبیت پلیمرز DNA به جای DNA، مونیتورینگ بلادرنگ روند سنتز و طول خواندن بالقوه چند هزار پایه‌ای DNA بود [43]. تقریباً 100 زیپتولیتری فراهم می‌کند که کاهش نویز پس زمینه، موازی سازی، و مونیتورینگ پلیمریزاسیون DNA تک مولکولی را امکان پذیر می‌کند. قالب DNA مجاز به انتشار در ZMV در حضور پرایمر و نوکلئوتیدها با برچسب‌های فلورسنت متصل به زنجیره فسفات است. پلیمرز DNA روی قالب DNA چفت می‌شود و شروع به ترکیب نوکلئوتیدهای مجزای برچسب گذاری شده می‌کند. تعیین توالی DNA با ثبت هر رویداد ترکیب نوکلئوتیدی دیکته شده توسط قالب دی ان ای انجام می‌شود. نوکلئوتیدهایی که در حال ترکیب هستند نسبت به نوکلئوتیدهایی که آزادانه انتشار می‌یابند برای مدت طولانی‌تری در حجم ردیابی نوری باقی می‌مانند، که جداسازی نویز پس زمینه تولید شده توسط غلظت بالای نوکلئوتیدهای برچسب گذاری شده از رویدادهای ترکیب نوکلئوتیدهای واقعی را تسهیل می‌کند. برچسب متصل به زنجیره فسفات در طول فرایند طویل سازی جدا می‌شود. پلیمرزها به صورت تصادفی در میان ZMW ها بی حرکت می‌شوند، که منجر می‌شود تنها یک سوم از آن‌ها توسط تک پلیمرز اشغال شوند. بقیه ZMW ها یا خالی و یا حاوی دو یا چند پلیمرز هستند، که به طور چشمگیری ظرفیت بالقوه آرایه ZMW را کاهش می‌دهد. توالی SMRT نرخ نرخ خطای بالایی دارد، عمدتاً به دلیل شکست در شناسایی تمام پیوستگی‌های منجر به جهش‌های رخنه‌ای یا حذفی است [42, 43].



شکل 2. (A) مقایسه تعداد مولکول‌های DNA مورد نیاز برای تولید یک پاسخ پایه در تعیین توالی اجماعی و تعیین توالی تک مولکولی. (B) شایع‌ترین نوع اشتباهات در تکنولوژی تعیین توالی.

توالی SMRT دارای چندین ویژگی جالب است که می‌تواند مورد سوء استفاده قرار گیرد. یک قالب دایره می‌تواند مورد استفاده قرار گیرد، که دستیابی به یک توالی اجماعی با کیفیت بالا را با اجازه دادن به پلیمرز برای گذرهای متعدد در امتداد قالب DNA یکسان ممکن می‌سازد [45]. طول خواندن فیزیکی می‌تواند با خاموش کردن لیزر مورد استفاده برای درخشان کردن ZMW ها در فواصل از پیش تعیین شده گسترش یابد، که subreads هایی از بخش‌های یکسان از ژنومیک در فرآیندی به نام توالی strobe تولید می‌کند [46]. یکی از ویژگی‌های واقعا جدید قابلیت ضبط اطلاعات جنبشی با مشاهده فعالیت یک آنزیم غیرفعال بصورت بلادرنگ است. این، ردیابی تغییرات شیمیایی در پایه‌ها مانند متیلاسیون [47]، و مشاهده سینتیک ترجمه با استفاده از tRNA های برچسب گذاری شده و بیحرکت سازی ریبوزوم به جای پلیمرز در ZMW را ممکن می‌سازد [48]. رویکرد تعیین توالی SMRT دارای پتانسیل بسیار زیادی با ویژگی‌های منحصر به فرد و اجرای واکنش تعیین توالی سریعی دارد، اما از نرخ خطای خام بیش از 10٪ رنج می‌برد، که عمدتا به دلیل جهش‌های رخنه‌ای یا حذفی، و خروجی نسبتا کم است که مانع استفاده از آن در بسیاری از کاربردها می‌شود. توالی تک مولکولی بلادرنگ (SMRT) متکی بر فرایند تثبیت پلیمرز DNA از طریق یک ارتباط بیوتین- استرپتاویدین در پایین یک نانو ساختار است که موجب حالت صفر (ZMW) نامیده می‌شود [44]. این ساختار مهم است، زیرا یک فضای مشاهدات نوری محدود

4.3. تعیین توالی نانوحفره

در سال 1996، دیمر و همکاران نشان دادند که تک رشته DNA (1.1-1.3 نانومتر؛ [49]) و یا مولکول‌های RNA می‌تواند از طریق یک کانال یونی در یک لایه دوتایی لیپیدی توسط میدان الکتریکی تحریک شود [50]. این مقاله استناد شده قبلاً اصول اساسی روش تعیین توالی نانوحفره را تعیین کرده است. برای تعیین توالی، DNA با مونیتورینگ یک جریان الکتریکی همچنانکه DNA از طریق یک نانوحفره (یک کانال با قطر 1-10 نانومتر) که در یک غشا عبور می‌کرد رمزگشایی می‌شد، که می‌توانست به صورت nanofabricated باشد یا توسط یک پروتئین مهندسی شده ایجاد شود. تعیین توالی نانوحفره به طور بالقوه می‌تواند آماده سازی نمونه حداقلی، بازخوانی الکتریکی یا فلورسنت و خواندن طول چند هزار پایه‌ای از مولکول‌های تکی DNA بصورت بلادرنگ را پیشنهاد دهد [42, 51].

قبل از اینکه تعیین توالی با نانوحفره‌ها بتواند به طور کامل استفاده قرار گیرد، چندین مشکل باید حل شود. نانوحفره‌ها بیش از حد طولانی هستند تا تفکیک تک پایه‌ای ارائه کنند همچنانکه چند نوکلئوتید در منفذ، در هر تغییری در جریان الکتریکی شرکت می‌کنند. علاوه بر این، سرعت بالای انتقال DNA از طریق نانوحفره‌ها اندازه گیری‌های الکتریکی را مشکل می‌کند، که تفسیر سیگنال شناسایی نوکلئوتید را دشوار می‌کند [42]. با این حال، راه‌حلهایی برای این مشکلات ارائه شده است. آی بی ام در حال توسعه یک ماتریس نانو حفره است، که شبیه ترانزیستوری با لایه‌های متناوب از فلز و ماده دی الکتریک است. شبیه سازی نشان می‌دهد که تعدیل جریان در این "ترانزیستور" می‌تواند سرعت انتقال DNA را کنترل کند [52]. با این حال، نیاز است این اثر بطور تجربی تایید شود و یک دستگاه پتانسیل هنوز برای دستیابی به وضوح تک پایه‌ای نیاز است [42, 52]. Genia یک شرکت تعیین توالی نانوحفره با استفاده از لایه دوگانه لیپیدی با تک نانوحفره متصل به سنسور روی یک مدار مجتمع است. هدف آنها رسیدن به یک تراشه حاوی حدوداً 1 میلیون از چنین سنسورهایی برای راه اندازی تجاری سیستمشان است که برای سال 2013 برنامه ریزی شده است. با این حال، هیچ گزارشی از غیر فروشندگان وجود ندارد که سیستم GENIA را آزمایش کرده باشند. یکی دیگر از بازیگران تجاری که برای رسیدن به تعیین توالی نانوحفره تلاش می‌کند Oxford Nanopore Technologies است. آنها در حال حاضر دو استراتژی برای تعیین توالی DNA دارند: تعیین توالی رشته‌ای و اگزونوکلاز. توالی اگزونوکلاز یک α

همولیزین اصلاح شده با یک اگزونوکلاز متصل است، که در داخل یک غشاء سینتتیک با مقاومت بالا الکترونیکی واقع شده است. اگزونوکلاز تک نوکلئوتید را می‌شکافد و آنها را در میان منافذ تغذیه می‌کند. هر تک نوکلئوتید می‌تواند از طریق سیگنال الکتریکی متمایزی تشخیص داده شود همچنانکه آنها بطور گذرا به مولکول سیکلودکسترین در هنگام عبور از میان منافذ متصل می‌شوند [53]. تعیین توالی رشته‌ای متکی بر پلیمر است که (با دیگر آنزیم اصلاح DNA) تک رشته‌ای را در داخل منافذ تغذیه می‌کند و سیگنال الکتریکی از سه را ردیابی می‌کند، برای مثال، پایه‌هایی که با یک منطقه خاص از منفذ پروتئین تعامل دارند.

به تازگی، Oxford Nanopore Technologies توانایی خود برای تعیین توالی کل 48000 پایه ژنوم لامبدا روی هر دو رشته سنس و آنتی‌سنس با استفاده از روش تعیین توالی رشته‌ای را اعلان کرد. تجاری سازی شیمی nanopore-based sensing بر روی یک پلتفرم با پایه الکترونیک، GridION، و ابزارهای قابل حمل یکبار مصرف برای سنجش الکترونیکی تک مولکول، Min-ION، برای 2012 برنامه ریزی شده است. Oxford Nanopore Technologies ادعای دقت رقابتی و طول خواندن فوق العاده طولانی داده‌های تک مولکولی را دارد، که ژنوم کامل انسان را در 15 دقیقه با استفاده از GridIONs های متعدد ارائه خواهد کرد. خیلی زود است بگوییم که آیا Oxford Nanopore Technologies می‌تواند به طور کامل بر روی این وعده بماند و آزمون اسید برای تکنیک تعیین توالی ارزیابی از آزمونگر اولیه خواهد بود. با این حال، اگر وعده دقت رقابتی و خواندن فوق العاده طولانی صادق باشد است باعث تغییر دیگر پارادایم‌ها در زمینه تعیین توالی می‌شود.

5. چشم انداز آینده

تکنولوژی‌های تعیین توالی پیشرفت‌های سریعی در سال‌های اخیر داشته‌اند و هیچ نشانه‌ای از کاهش پیشرفت وجود ندارد. این مساله عدم تعادلی بین تعیین توالی برای هر se و دیگر روش‌های درگیر، در هر دو مراحل بالادستی و پایین دستی، به عنوان مثال آماده سازی نمونه و تجزیه و تحلیل داده‌ها و ذخیره‌سازی ایجاد کرده است. بارکدگذاری نمونه‌ها به همراه اتوماسیون مراحل پیش از تعیین توالی پتانسیل این را دارد که در بسیاری از مسائل مربوط به فرآیند بالادستی استفاده شود. چالش واقعی در مراحل پایین دستی تجزیه و تحلیل داده‌های توالی، حمل و ذخیره‌سازی است. ذخیره سازی طولانی مدت داده‌های توالی و تجزیه و تحلیل‌های بعدی دست و

پا گیر و گران است، بنابراین، در آینده، ممکن است تجدید تعیین توالی نمونه‌ها ارزان‌تر و آسان‌تر باشد چنانچه اطلاعاتی مورد نیاز باشد. البته، چالش‌های متعددی وجود دارد که در واکنش‌های تعیین توالی واقعی باقی می‌ماند. دو تا از ضروری‌ترین موارد، طول خواندن و کیفیت خواندن دنباله است که وضوح تکرارهای مرسوم و نواحی تکثیر شده را ممکن می‌سازد. در حال حاضر، پلتفرم‌های تعیین توالی موازی فشرده و تک مولکولی مکمل هستند، و هیچ تکنولوژی بتنهایی برای تمام کاربردها بهترین گزینه نیست. بنابراین باید دید کدام تکنولوژی منجر به تغییراتی در پارادایم تعیین توالی خواهد شد.

توالی به طور سنتی برای مراکز تحقیقاتی بزرگ با بودجه بسیار خوب محفوظ است، اما هزینه برای هر پایه توالی در طول پایان HGP شروع به کاهش کرده است. سرعت کاهش هزینه‌ها حتی با راه‌اندازی و متعاقب آن تکامل ابزارهای تعیین توالی موازی فشرده بنظر می‌رسد در سال‌های اخیر بیشتر شتاب گرفته است.

دموکراتیزه کردن تعیین توالی از افزایش مداوم و قابل ملاحظه تعداد کاربردها و تحلیل‌هایی که شامل داده‌های توالی است مشهود است. امروزه تعیین توالی در عرصه‌های گوناگون، مانند پزشکی قانونی، کشاورزی، تولید سوخت‌های زیستی، دیرین‌شناسی، اهلی، و داروسازی کاربرد دارد. این انقلاب در حال وقوع وعده آوردن تعیین توالی به تقریباً هر جنبه‌ای از زندگی را می‌دهد.

References

- [1] Maxam, A. M., Gilbert, W., A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA* 1977, 74, 560–564.
- [2] Sanger, F., Nicklen, S., Coulson, A. R., DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 1977, 74, 5463–5467.
- [3] Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P. et al., Fluorescence detection in automated DNA sequence analysis. *Nature* 1986, 321, 674–679.
- [4] Prober, J. M., Trainor, G. L., Dam, R. J., Hobbs, F. W. et al., A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* 1987, 238, 336–341.
- [5] Cohen, A. S., Najarian, D. R., Paulus, A., Guttman, A. et al., Rapid separation and purification of oligonucleotides by high-performance capillary gel electrophoresis. *Proc. Natl. Acad. Sci. USA* 1988, 85, 9660–9663.
- [6] Luckey, J. A., Drossman, H., Kostichka, A. J., Mead, D. A. et al., High speed DNA sequencing by capillary electrophoresis. *Nucleic Acids Res.* 1990, 18, 4417–4421.
- [7] Karger, B. L., Guttman, A., DNA sequencing by CE. *Electrophoresis* 2009, 30 Suppl 1, S196–202.
- [8] Melamede, R. J., Automatable process for sequencing nucleotide. US Patent 4863849, 1985.
- [9] Nyren, P., The history of pyrosequencing. *Methods Mol. Biol.* 2007, 373, 1–14.
- [10] Ronaghi, M., Uhlen, M., Nyren, P., A sequencing method based on real-time pyrophosphate. *Science* 1998, 281, 363–365.
- [11] Margulies, M., Egholm, M., Altman, W. E., Attiya, S. et al., Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 2005, 437, 376–380.
- [12] Rothberg, J. M., Leamon, J. H., The development and impact of 454 sequencing. *Nat. Biotechnol.* 2008, 26, 1117–1124.
- [13] Wheeler, D. A., Srinivasan, M., Egholm, M., Shen, Y. et al., The complete genome of an individual by massively parallel DNA sequencing. *Nature* 2008, 452, 872–876.
- [14] Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C. et al., Initial sequencing and analysis of the human genome. *Nature* 2001, 409, 860–921.
- [15] Green, R. E., Krause, J., Briggs, A. W., Maricic, T. et al., A draft sequence of the Neandertal genome. *Science* 2010, 328, 710–722.
- [16] Bennett, S. T., Barnes, C., Cox, A., Davies, L., Brown, C., Toward the \$1000 human genome. *Pharmacogenomics* 2005, 6, 373–382.
- [17] Balasubramanian, S., Bentley, D., Polynucleotide arrays and their use in sequencing. Patent WO 01/157248, 2001.
- [18] Shendure, J., Ji, H., Next-generation DNA sequencing. *Nat. Biotechnol.* 2008, 26, 1135–1145.
- [19] Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P. et al., Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 2008, 456, 53–59.
- [20] Bentley, D. R., Whole-genome re-sequencing. *Curr. Opin. Genet. Dev.* 2006, 16, 545–552.
- [21] Sanger, F., Coulson, A. R., A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.* 1975, 94, 441–448.
- [22] Li, R., Zhu, H., Ruan, J., Qian, W. et al., De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* 2010, 20, 265–272.
- [23] Li, R., Fan, W., Tian, G., Zhu, H. et al., The sequence and de novo assembly of the giant panda genome. *Nature* 2010, 463, 311–317.
- [24] Lehman, I. R., DNA Ligase: Structure, mechanism, and function. *Science* 1974, 186, 790–797.
- [25] Shendure, J., Porreca, G. J., Reppas, N. B., Lin, X. et al., Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* 2005, 309, 1728–1732.
- [26] Metzker, M. L., Sequencing technologies – the next generation. *Nat. Rev. Genet.* 2010, 11, 31–46.
- [27] McKernan, K. J., Peckham, H. E., Costa, G. L., McLaughlin, S. F. et al., Sequence and structural variation in a human genome uncovered by short-read, massively parallel ligation sequencing using two-base encoding. *Genome Res.* 2009, 19, 1527–1541.
- [28] Valouev, A., Ichikawa, J., Tonthat, T., Stuart, J. et al., A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Res.* 2008, 18, 1051–1063.

- [29] Bao, S., Jiang, R., Kwan, W., Wang, B. et al., Evaluation of next-generation sequencing software in mapping and assembly. *J. Hum. Genet.* 2011, 56, 406–414.
- [30] Drmanac, R., Sparks, A. B., Callow, M. J., Halpern, A. L. et al., Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science* 2010, 327, 78–81.
- [31] Fire, A., Xu, S. Q., Rolling replication of short DNA circles. *Proc. Natl. Acad. Sci. USA* 1995, 92, 4641–4645.
- [32] Pourmand, N., Karhanek, M., Persson, H. H. J., Webb, C. D. et al., Direct electrical detection of DNA synthesis. *Proc. Natl. Acad. Sci. USA* 2006, 103, 6466–6470.
- [33] Rothberg, J. M., Hinz, W., Rearick, T. M., Schultz, J. et al., An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 475, 348–352.
- [34] Toumazou, C., Premanode, B., Shepherd, L., Signal processing circuit comprising ion sensitive field effect transistor and method of monitoring a property of a fluid, *DNA Electronics*, US Patent 20080265985, 2008.
- [35] Mellmann, A., Harmsen, D., Cummings, C. A., Zentz, E. B. et al., Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology. *PLoS ONE* 2011, 6, e22751.
- [36] Rohde, H., Qin, J., Cui, Y., Li, D. et al., Open-source genomic analysis of shiga-toxin-producing *E. coli* O104:H4. *N. Engl. J. Med.* 2011, 365, 718–724.
- [37] Efcavitch, J. W., Thompson, J. F., Single-molecule DNA analysis. *Annu. Rev. Anal. Chem. (Palo Alto Calif)* 2010, 3, 109–128.
- [38] Harris, T. D., Buzby, P. R., Babcock, H., Beer, E. et al., Single-molecule DNA sequencing of a viral genome. *Science* 2008, 320, 106–109.
- [39] Bowers, J., Mitchell, J., Beer, E., Buzby, P. R. et al., Virtual terminator nucleotides for next-generation DNA sequencing. *Nat. Methods* 2009, 6, 593–595.
- [40] Pushkarev, D., Neff, N. F., Quake, S. R., Single-molecule sequencing of an individual human genome. *Nat. Biotechnol.* 2009, 27, 847–850.
- [41] Ozsolak, F., Platt, A. R., Jones, D. R., Reifengerger, J. G. et al., Direct RNA sequencing. *Nature* 2009, 461, 814–818.
- [42] Schadt, E. E., Turner, S., Kasarskis, A., A window into third-generation sequencing. *Hum. Mol. Genet.* 2010, 19, R227–240.
- [43] Eid, J., Fehr, A., Gray, J., Luong, K. et al., Real-time DNA sequencing from single polymerase molecules. *Science* 2009, 323, 133–138.
- [44] Korlach, J., Bjornson, K. P., Chaudhuri, B. P., Cicero, R. L. et al., Realtime DNA sequencing from single polymerase molecules. *Methods Enzymol.* 2010, 472, 431–455.
- [45] Travers, K. J., Chin, C.-S., Rank, D. R., Eid, J. S., Turner, S. W., A flexible and efficient template format for circular consensus sequencing and SNP detection. *Nucleic Acids Res.* 2010, 38, e159.
- [46] Lo, C., Bashir, A., Bansal, V., Bafna, V., Strobe sequence design for haplotype assembly. *BMC Bioinformatics* 2011, 12, S24.
- [47] Flusberg, B. A., Webster, D. R., Lee, J. H., Travers, K. J. et al., Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat. Methods* 2010, 7, 461–465.
- [48] Uemura, S., Aitken, C. E., Korlach, J., Flusberg, B. A. et al., Real-time tRNA transit on single translating ribosomes at codon resolution. *Nature* 2010, 464, 1012–1017.
- [49] Mandelkern, M., Elias, J. G., Eden, D., Crothers, D. M., The dimensions of DNA in solution. *J. Mol. Biol.* 1981, 152, 153–161.
- [50] Kasianowicz, J. J., Brandin, E., Branton, D., Deamer, D. W., Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Acad. Sci. USA* 1996, 93, 13770–13773.
- [51] Branton, D., Deamer, D. W., Marziali, A., Bayley, H. et al., The potential and challenges of nanopore sequencing. *Nat. Biotechnol.* 2008, 26, 1146–1153.
- [52] Luan, B., Peng, H., Polonsky, S., Rossmagel, S. et al., Base-by-base ratcheting of single stranded DNA through a solid-state nanopore. *Phys. Rev. Lett.* 2010, 104, 238103.
- [53] Clarke, J., Wu, H.-C., Jayasinghe, L., Patel, A. et al., Continuous base identification for single-molecule nanopore DNA sequencing. *Nat. Nanotechnol.* 2009, 4, 265–270.