

# เทคโนโลยีการวิเคราะห์ลำดับเบสยุคที่สองและสาม

## Second and Third Generation Sequencing Technologies

กิตติพัฒน์ อุโฆษกิจ\*, มาริสา รักสุขสมบัติ และสุภาภรณ์ ขานโบ

ภาควิชาเทคโนโลยีชีวภาพ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์  
ศูนย์รังสิต ตำบลคลองหนึ่ง อำเภอคลองหลวง จังหวัดปทุมธานี 12120

Kittipat Ukoskit\*, Marisa Raksuksombat and Supaporn Khanbo

Department of Biotechnology, Faculty of Science and Technology, Thammasat University,  
Rangsit Centre, Khlong Nueng, Khlong Luang, Pathum Thani 12120

### บทคัดย่อ

การหาลำดับเบสโดยวิธีของแซงเกอร์นับเป็นเทคโนโลยีการวิเคราะห์ลำดับเบสยุคที่หนึ่งและได้ถึงจุดสิ้นสุดเมื่อเกิดเทคโนโลยีการวิเคราะห์ลำดับเบสยุคใหม่ ซึ่งเป็นเทคโนโลยีการวิเคราะห์ลำดับเบสยุคที่สอง (second generation sequencing technologies, SGSTs) และสาม (third generation sequencing technologies, TGSTs) ความแตกต่างของเทคโนโลยีการวิเคราะห์ลำดับเบสยุคที่หนึ่งจาก SGSTs และ TGSTs คือความสามารถของเทคโนโลยีในการวิเคราะห์ลำดับเบสเป็นจำนวนมาก ๆ ในเวลาเดียวกัน SGSTs ได้รับการยอมรับและมีการใช้อย่างแพร่หลายในช่วงหลายปีที่ผ่านมา รูปแบบหลักของเทคโนโลยี SGSTs เช่น 454/Roche, Solexa, MiSeq, HiSeq/Illumina, SOLiD/Life Technology และ Ion Torrent/Life Technology รูปแบบของเทคโนโลยีเหล่านี้แตกต่างกันโดยหลักการของการวิเคราะห์ลำดับเบส จำนวนและรูปแบบข้อมูลลำดับเบสที่ได้ โดยสิ่งที่เหมือนกันของทุกเทคโนโลยี SGSTs คือการเพิ่มปริมาณดีเอ็นเอต้นแบบเพื่อใช้ในการวิเคราะห์ลำดับเบส เนื่องจากเทคโนโลยีไม่มีความไวพอในการวิเคราะห์ลำดับเบสโมเลกุลเดี่ยว SGSTs สามารถให้ข้อมูลสายดีเอ็นเอที่มีความยาวตั้งแต่ 25-800 คู่เบส จำนวนหลายล้านถึงหลายพันล้านเส้น แล้วแต่เทคโนโลยี ใช้เวลาในการวิเคราะห์เพียงไม่กี่วัน และมีค่าใช้จ่ายต่ำเมื่อเทียบกับเทคโนโลยีการวิเคราะห์ลำดับเบสยุคที่หนึ่ง แม้ว่า SGSTs จะประสบความสำเร็จทางการค้า แต่ก็ยังมีความเร็วไม่มากพอ ให้ข้อมูลลำดับเบสที่ครอบคลุมไม่เพียงพอทั้งจีโนม และอาจให้ข้อมูลลำดับเบสที่มีความลำเอียงเนื่องจากต้องมีการเพิ่มปริมาณดีเอ็นเอต้นแบบ TGSTs คิดค้นขึ้นเพื่อใช้ในการวิเคราะห์โมเลกุลเดี่ยวของดีเอ็นเอหรืออาร์เอ็นเอโดยซึ่งเป็นโมเลกุลธรรมชาติ TGSTs มีศักยภาพมากขึ้น สามารถวิเคราะห์ลำดับเบสดีเอ็นเอได้สายยาวมากขึ้น มีค่าใช้จ่ายลดลง ใช้ระยะเวลาในการวิเคราะห์การลดลง และสามารถลดความลำเอียงของการวิเคราะห์ลำดับเบสอันเกิดจากการเพิ่มดีเอ็นเอต้นแบบ นอกจากนี้ยังสามารถประยุกต์ใช้ในงานอื่นนอกเหนือจากการวิเคราะห์ลำดับเบสดีเอ็นเอได้ ในบทความนี้เป็นบทความโดยสังเขปถึงสถานะและรูปแบบของ SGSTs และการเกิดขึ้นของ TGSTs รวมถึงการประยุกต์ใช้เทคโนโลยีการวิเคราะห์ลำดับเบสยุคใหม่นี้

คำสำคัญ : next generation sequencing (NGS); second generation sequencing technologies (SGSTs); third generation sequencing technologies (TGSTs)

\*ผู้รับผิดชอบบทความ : ku@tu.ac.th

## Abstract

The era of first generation sequencing dominated by Sanger sequencing quickly came to an end with the introduction of next generation sequencing (NGS). Second and third generation DNA sequencing technologies are emerged as high-throughput approaches to DNA sequencing using the concept of massively parallel processing. The second generation sequencing technologies (SGSTs) have become widely available over the past several years. The 454/Roche, Solexa/Illumina, SOLiD/Life Technology (ABI) and Ion Torrent are currently the major platforms of SGSTs that differ in sequencing principles, number of read outputs and sequence read formats. The common principle of SGSTs is their need for clonally amplified molecules for sequencing signal detection because of their lack of single molecule sensitivity. These SGSTs generate hundreds of thousands to billions of 25-800 nucleotide-long reads within days in a low-cost manner compared to first generation sequencing. Although the scientific and commercial success of SGSTs, the goal of rapid, comprehensive and unbiased sequencing of nucleic acids has not been achieved. Third generation sequencing technologies (TGSTs) have emerged to analyze natural DNA and RNA molecules in a massively parallel manner with the single molecule DNA sequencing and direct RNA sequencing. TGSTs have potential to increase sequencing throughput and read lengths, decrease costs, run times and error rates, eliminate biases inherent in SGSTs, and offer capabilities beyond DNA sequencing. In this review, we briefly summarize the current state of SGSTs and the emerging TGSTs and outline the broad range of applications for NGS technologies.

**Keywords:** next generation sequencing (NGS); second generation sequencing technologies (SGSTs); third generation sequencing technologies (TGSTs)

## 1. บทนำ

ปัจจุบันเทคโนโลยีการหาลำดับเบสดีเอ็นเอมีความก้าวหน้าอย่างมาก มีการพัฒนาเทคนิคใหม่ ๆ ทำให้สามารถหาลำดับเบสได้รวดเร็วขึ้น และมีต้นทุนค่าใช้จ่ายที่ต่ำลงต่อจำนวนข้อมูลของลำดับเบสวิเคราะห์ได้ ซึ่งแตกต่างจากการหาลำดับเบสดีเอ็นเอโดยวิธีมาตรฐานหรือการหาลำดับเบสโดยวิธีแซงเกอร์ (Sanger sequencing) ที่การหาลำดับเบสทั้งจีโนมของสิ่งมีชีวิตเป็นเรื่องที่ยาก ใช้เวลานาน และมีค่าใช้จ่ายสูง การหาลำดับเบสดีเอ็นเอโดยวิธีมาตรฐานถูกพัฒนาขึ้นในปี ค.ศ. 1977 โดย Frederick Sanger

[1] เทคนิคนี้อาศัยการใช้ dideoxynucleotide triphosphates (ddNTPs) เป็นตัวหยุดการสังเคราะห์สายดีเอ็นเอ โดยแบ่งปฏิกิริยาการสังเคราะห์ดีเอ็นเอเป็นสี่หลอดในแต่ละปฏิกิริยาจะมีไพรเมอร์ซึ่งจะเข้าจับกับดีเอ็นเอต้นแบบ และเอนไซม์ดีเอ็นเอพอลิเมอเรส (DNA polymerase) ทำหน้าที่สังเคราะห์สายดีเอ็นเอโดยใช้ deoxynucleotide triphosphates (dNTP) ที่เติมลงในปฏิกิริยาเพื่อสังเคราะห์สายดีเอ็นเอ ในแต่ละหลอดของปฏิกิริยาการสังเคราะห์ดีเอ็นเอจะใส่ ddNTP คนละชนิด (ddATP, ddCTP, ddGTP หรือ ddTTP) ปฏิกิริยาการสังเคราะห์ดีเอ็นเอจะหยุดเมื่อมี

ddNTP เข้ามาจับ ทำให้แต่ละปฏิกิริยาจะมีชิ้นดีเอ็นเอหลากหลายขนาดรวมกันอยู่ ผลผลิตของทั้งสี่ปฏิกิริยาถูกนำมาวิเคราะห์ขนาดด้วยเจลพอลิอะครีลาไมด์ (polyacrylamide gel) และใช้วิธีอโตเรดิโอกราฟ (autoradiograph) แสดงแถบดีเอ็นเอ ต่อมาเทคนิคนี้ได้มีการปรับปรุงให้มีประสิทธิภาพมากขึ้น โดยมีการใช้สีย้อมฟลูออเรสเซนส์ (fluorescence dye) [2] แทนการใช้สารกัมมันตรังสีในวิธีอโตเรดิโอกราฟ และรวมปฏิกิริยาทั้งสี่ทำในหลอดเดียวกัน นอกจากนี้มีการใช้อิเล็กโตรโฟรีซิสแบบแคปิลลารี (capillary electrophoresis) แทนที่การแยกชิ้นดีเอ็นเอโดยพอลิอะครีลาไมด์เจล วิธีการนี้นับได้ว่าเป็นเทคโนโลยีการวิเคราะห์ลำดับเบสยุคที่หนึ่ง (first generation sequencing) วิธีแซงเกอร์สามารถวิเคราะห์ลำดับเบสได้ในช่วงประมาณ 500 ถึง 800 คู่เบส เทคนิคนี้ถูกนำมาใช้ในงานวิจัยต่าง ๆ มากมายรวมถึงโครงการหาลำดับเบสจีโนมมนุษย์ (human genome project) เนื่องจากข้อจำกัดของวิธีแซงเกอร์ที่อ่านลำดับเบสได้ไม่เกิน 800 คู่เบส จึงได้มีการพัฒนาวิธีการวิเคราะห์ลำดับเบสดีเอ็นเอที่มีขนาดใหญ่ที่เรียกว่าการวิเคราะห์ลำดับเบสแบบช็อตกัน (shotgun sequencing) โดยตัดชิ้นดีเอ็นเอที่มีขนาดใหญ่ให้มีขนาดเล็กลงและให้มีส่วนซ้อนทับกัน (overlapping) ของชิ้นดีเอ็นเอเพื่อสามารถนำข้อมูลลำดับเบสมาเชื่อมกันได้เป็นข้อมูลลำดับเบสของดีเอ็นเอที่มีขนาดใหญ่ที่วิเคราะห์

ปัจจุบันเทคโนโลยีการวิเคราะห์ลำดับเบสได้ก้าวเข้าสู่เทคโนโลยีการวิเคราะห์ลำดับเบสยุคใหม่ (next-generation sequencing, NGS) ซึ่งในบทความนี้จะหมายถึงเทคโนโลยีการวิเคราะห์ลำดับเบสยุคที่สองและยุคที่สามเทคโนโลยีการวิเคราะห์ลำดับเบสยุคที่สอง (second generation sequencing technologies, SGSTs) สามารถวิเคราะห์ลำดับเบสของดีเอ็นเอได้พร้อม ๆ กันจำนวนมาก (massively

parallel sequencing) จากคลังดีเอ็นเอ (DNA library) ต้นแบบที่เตรียมโดยปฏิกิริยาพีซีอาร์ ทำให้ได้ข้อมูลลำดับเบสจำนวนมากมหาศาล ในเวลาอันรวดเร็ว SGSTs ที่คิดค้นขึ้นมาในแต่ละรูปแบบประสบความสำเร็จเป็นอย่างดีในทางการค้า และมีประโยชน์อย่างมากกับงานวิจัยทางจีโนมและพันธุศาสตร์ของมนุษย์ สัตว์ พืช และจุลินทรีย์ ทำให้ได้ความรู้และข้อมูลที่สามารถนำไปประยุกต์ใช้ในการแพทย์ การเกษตร และอุตสาหกรรม

ถึงแม้ว่า SGSTs จะก้าวหน้าอย่างรวดเร็วสามารถผลิตข้อมูลลำดับเบสได้จำนวนมาก โดยมีค่าใช้จ่ายต่อเบสต่ำ ปัจจุบันได้มีเทคโนโลยีการวิเคราะห์ลำดับเบสยุคที่สาม (third generation sequencing technologies, TGSTs) เป็นเทคโนโลยีการวิเคราะห์ลำดับเบสที่ไม่จำเป็นต้องมีการสร้างคลังดีเอ็นเอต้นแบบโดยการทำปฏิกิริยาพีซีอาร์ แต่เป็นการวิเคราะห์ลำดับเบสจากดีเอ็นเอโมเลกุลเดี่ยวแต่ละโมเลกุล (single molecule sequencing) เป็นโมเลกุลดีเอ็นเอธรรมชาติที่ไม่ได้เกิดจากการสังเคราะห์ในหลอดทดลอง ดังนั้นข้อมูลลำดับเบสที่ได้จึงไม่มีความคลาดเคลื่อนอันเกิดจากการเพิ่มปริมาณดีเอ็นเอในปฏิกิริยาพีซีอาร์ นอกจากนี้ข้อมูลลำดับเบสที่ได้มีความยาวของข้อมูลลำดับเบสเพิ่มขึ้นจาก SGSTs อย่างมาก อาจมีความยาวได้หลายสิบกิโลเบสของข้อมูลดีเอ็นเอหนึ่งสาย นอกจากนี้เวลาที่ใช้ในการวิเคราะห์ก็อาจลดลงเหลือแค่ไม่กี่วินาทีหรือชั่วโมงหรือแม้แต่นาที TGSTs นี้มีข้อได้เปรียบกว่า SGSTs หลายประการเช่น ได้ข้อมูลจำนวนมากกว่า สามารถวิเคราะห์ลำดับเบสจีโนมของสิ่งมีชีวิตขนาดเล็กในเวลาอย่างรวดเร็วและมีจำนวนซ้ำของข้อมูล (coverage) เป็นจำนวนมาก ได้ข้อมูลลำดับเบสสายยาวมากจึงทำให้สามารถต่อข้อมูลลำดับเบสเป็นดีเอ็นเอสายยาวโดยไม่ต้องเทียบข้อมูลจีโนมอ้างอิง (de novo assembly)

ใช้ดีเอ็นเอเริ่มต้นเพียงเล็กน้อย (โดยทฤษฎีแล้วดีเอ็นเอหนึ่งโมเลกุลเท่านั้นที่ใช้สำหรับการวิเคราะห์ลำดับเบสสำหรับ TGSs) มีค่าใช้จ่ายในการวิเคราะห์ลำดับเบสที่ค่อนข้างต่ำ การวิเคราะห์จีโนมของมนุษย์ที่มีจำนวนซ้ำของข้อมูลมากอาจใช้เงินเพียงสามหมื่นบาทเท่านั้น

## 2. เทคโนโลยีการวิเคราะห์ลำดับเบสยุคที่สอง

ปัจจุบัน SGSTs ถูกพัฒนาขึ้นมาอย่างรวดเร็ว โดยมีรูปแบบที่นิยมใช้อยู่ 4 รูปแบบ ได้แก่ 454/Roche, Solexa, MiSeq, HiSeq/Illumina,

SOLiD/Life Technology และ Ion Torrent/Life Technology ซึ่งแต่ละรูปแบบก็เหมาะต่อการนำไปใช้ในงานวิจัยที่แตกต่างกัน โดยแต่ละรูปแบบจะแตกต่างกันในเรื่องการเตรียมคลัสต์ดีเอ็นเอต้นแบบ การเพิ่มจำนวนดีเอ็นเอสำหรับวิเคราะห์ลำดับเบส วิธีการวิเคราะห์ลำดับเบส ความแตกต่างของข้อมูลความยาวลำดับเบสที่ได้ และระดับความถูกต้องของข้อมูล ดังแสดงเปรียบเทียบในตารางที่ 1 โดยแต่ละรูปแบบมีรายละเอียดการทำงานดังนี้

ตารางที่ 1 เปรียบเทียบรูปแบบของเทคโนโลยีการวิเคราะห์ลำดับเบสยุคที่สอง

| Platform                | Read length (bp) | Max. number of reads/run | Sequencing output/run | Run time   |
|-------------------------|------------------|--------------------------|-----------------------|------------|
| Roche 454 GS FLX        | 400-500          | 1 × 10 <sup>6</sup>      | ≤500 Mb               | 10 h       |
| Roche 454 GS FLX+       | 600-800          | 1 × 10 <sup>6</sup>      | ≤700 Mb               | 23 h       |
| Roche 454 GS Junior     | 400-450          | 1 × 10 <sup>5</sup>      | ~35 Mb                | 10 h       |
| Illumina HiSeq 2000     | 100-200          | 6 × 10 <sup>9</sup>      | ≤540-600 Gb           | 11 d       |
| Illumina HiSeq 1000     | 100-200          | 3 × 10 <sup>9</sup>      | ≤270-300 Gb           | 8.5 d      |
| Illumina GAIIx          | 50-75            | 6.4 × 10 <sup>8</sup>    | ≤95 Gb                | 7.5-14.5 d |
| Illumina MiSeq          | 100-150          | 7 × 10 <sup>6</sup>      | ≤1-2 Gb               | 19-27 h    |
| AB SOLiD 5500 system    | 35-75            | 2.4 × 10 <sup>9</sup>    | ~100 Gb               | 4 d        |
| AB SOLiD 5500 xl system | 35-75            | 6 × 10 <sup>9</sup>      | ~250 Gb               | 7-8 d      |
| Ion Torrent -314 chip   | 200-400          | 0.5 × 10 <sup>6</sup>    | ≤100 Mb               | 2.3-3.7 h  |
| Ion Torrent -316 chip   | 200-400          | 3 × 10 <sup>6</sup>      | ≤1 Gb                 | 3.0-4.9 h  |
| Ion Torrent -318 chip   | 200-400          | 5.5 × 10 <sup>6</sup>    | ≤2 Gb                 | 4.4-7.3 h  |
| Ion Torrent -PI chip    | 200              | 80 × 10 <sup>6</sup>     | ≤10 Gb                | 2-4 h      |

### 2.1 454/Roche genome sequencer

เทคโนโลยีการวิเคราะห์ลำดับเบสด้วยเครื่อง 454/Roche genome sequencer หรือที่รู้จักกันในชื่อเทคนิคการวิเคราะห์ลำดับเบสไฟโร (pyrosequencing) เป็นรูปแบบแรกของการวิเคราะห์ลำดับ

เบสยุคใหม่ที่ถูกนำออกมาใช้ในเชิงพาณิชย์ [3] เป็นวิธีการที่พัฒนาขึ้นในปี ค.ศ. 1996 โดย Nyrén และคณะ [4] หลักการวิเคราะห์ลำดับเบสด้วยวิธีการนี้อาศัยการทำงานของเอนไซม์ดีเอ็นเอพอลิเมอเรสในการสังเคราะห์ดีเอ็นเอสายคู่จากดีเอ็นเอต้นแบบสายเดียว

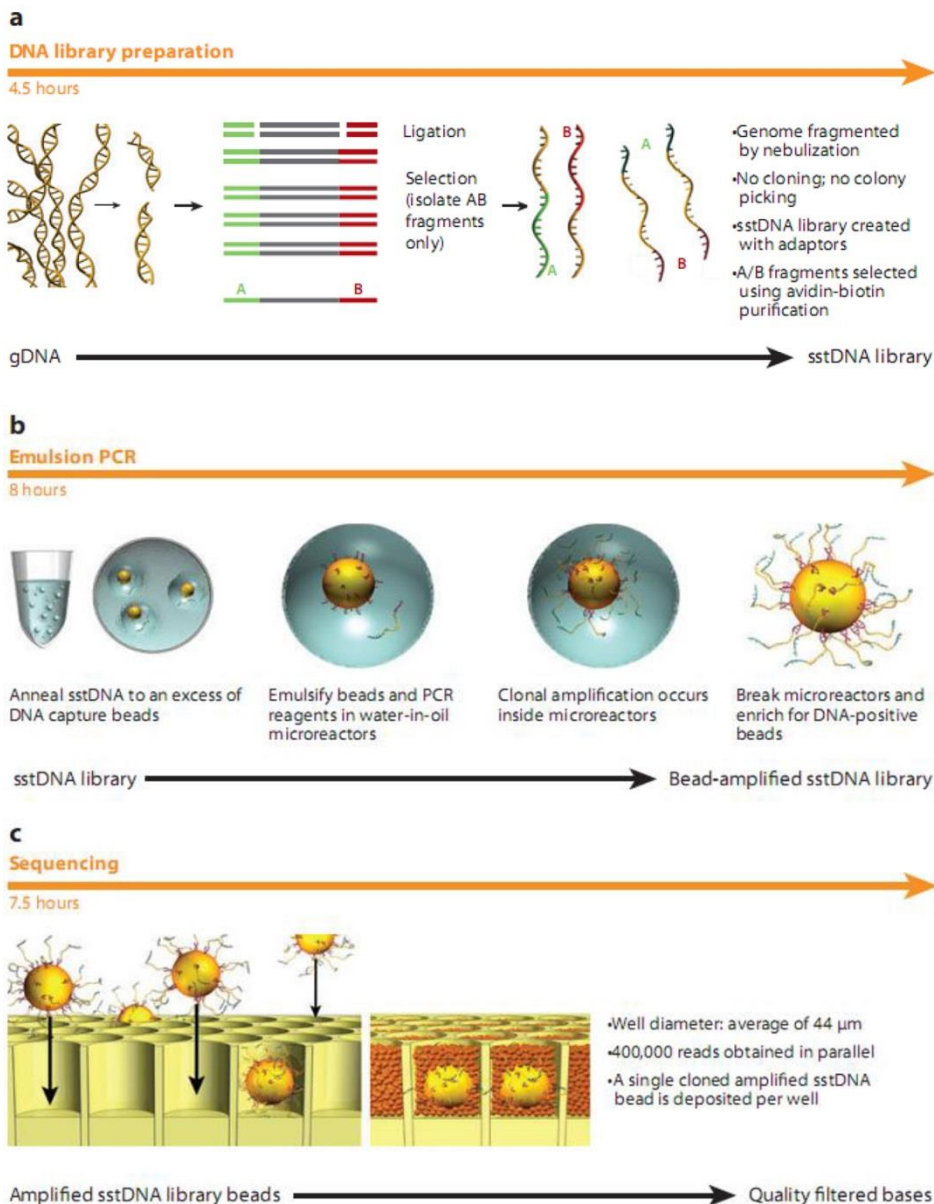
โดยในแต่ละรอบของการทำปฏิกิริยาจะใช้ dNTP เพียงชนิดเดียว การเข้าสู่ของนิวคลีโอไทด์กับดีเอ็นเอสายคู่สมเกิดขึ้นพร้อมกับการปลดปล่อยสารประกอบไพโรฟอสเฟต (pyrophosphate, PPi) ซึ่งจะถูกเปลี่ยนเป็นสัญญาณแสงจากปฏิกิริยาเคมี (chemiluminescence) โดยอาศัยการทำงานของเอนไซม์ 3 ชนิด คือ เอทีพีซัลฟูไรเลส (ATP sulfurylase) ลูซิเฟอเรส (luciferase) และอะไพเรส (apyrase) และสัญญาณแสงที่เกิดขึ้นจะถูกตรวจจับได้ด้วยกล้องบันทึกภาพตามจังหวะแสงของการเข้าเบสคู่สม

การเตรียมคลังดีเอ็นเอต้นแบบ และการเพิ่มปริมาณดีเอ็นเอสำหรับการวิเคราะห์ลำดับเบสด้วยเครื่อง 454/Roche genome sequencer เริ่มต้นจากการตัดดีเอ็นเอจีโนมแบบสุ่ม ให้เป็นสายสั้นประมาณ 400 ถึง 600 คู่เบส แล้วเชื่อมต่ออะแดปเตอร์ที่ต่างกัน (adapter A และ adapter B) ที่ปลายทั้งสองข้างของชิ้นดีเอ็นเอ จากนั้นชิ้นดีเอ็นเอที่เชื่อมต่ออะแดปเตอร์ที่ปลายทั้งสองข้างจะสามารถเข้าจับกับเม็ดลูกปัดซึ่งถูกเคลือบด้วยนิวคลีโอไทด์สายสั้น ๆ ที่มีเบสคู่สมกับอะแดปเตอร์ B ของชิ้นดีเอ็นเอ โดยสภาพแวดล้อมของปฏิกิริยาจะถูกปรับเพื่อให้หนึ่งเม็ดลูกปัดจับได้กับสายดีเอ็นเอเพียงหนึ่งโมเลกุลเท่านั้น ท้ายที่สุดจะได้ดีเอ็นเอสายเดี่ยว ซึ่งมีส่วนของอะแดปเตอร์ A ที่ปลาย 5' และ อะแดปเตอร์ B ที่ปลาย 3' ตรึงอยู่บนเม็ดลูกปัด หลังจากนั้นโมเลกุลดีเอ็นเอจะถูกเพิ่มปริมาณในหยดน้ำมันที่ประกอบด้วยสารสำหรับการทำปฏิกิริยาฟิซาร์อยู่ภายใน (oil aqueous emulsion) เรียกการเพิ่มปริมาณดีเอ็นเอแบบนี้ว่าอิมัลชันฟิซาร์ (emulsion PCR หรือ emulsion-based clonal amplification) เมื่อสิ้นสุดปฏิกิริยาฟิซาร์เม็ดลูกปัดในหยดน้ำมันจะมีดีเอ็นเอต้นแบบที่เหมือนกันเป็นล้านโมเลกุลตรึงอยู่

การวิเคราะห์ลำดับเบสไฟโร (รูปที่ 1) ใช้

หลักการของการสังเคราะห์สายดีเอ็นเอในการหาลำดับเบส (sequencing by synthesis) โดยอาศัยการทำงานของเอนไซม์ดีเอ็นเอพอลิเมอร์เรสในการสังเคราะห์ดีเอ็นเอสายใหม่จากดีเอ็นเอต้นแบบสายเดี่ยวที่ติดอยู่บนเม็ดลูกปัด ขั้นตอนการวิเคราะห์ลำดับเบสเริ่มจากการนำเม็ดลูกปัดที่ถูกเพิ่มปริมาณดีเอ็นเอต้นแบบด้วยวิธี อิมัลชันฟิซาร์ใส่ลงแผ่นที่มีหลุมขนาดเล็กเรียกว่า PicoTiterPlate™ (PTP) ในหนึ่งแผ่นประกอบด้วยหลุมประมาณสองล้านหลุม โดยแต่ละหลุมถูกออกแบบให้มีขนาดพอดีกับหนึ่งเม็ดลูกปัด จากนั้นกระบวนการวิเคราะห์ลำดับเบสในเครื่อง 454/Roche จะมีการเติมเอนไซม์และสารต่าง ๆ สำหรับการวิเคราะห์ลำดับเบสไฟโร รวมทั้งเบส A, T, C และ G แบบอัทโนมิติ โดยแต่ละรอบของปฏิกิริยาจะเติม dNTP ลงไปเพียงชนิดเดียว ขณะเกิดการสังเคราะห์ดีเอ็นเอมีการต่อสายของเบสคู่สมโดยเอนไซม์ดีเอ็นเอพอลิเมอร์เรส ปฏิกิริยาจะเกิดการปลดปล่อยสารประกอบไพโรฟอสเฟต PPi เอนไซม์เอทีพีซัลฟูไรเลส จะทำหน้าที่เปลี่ยนสารประกอบไพโรฟอสเฟต PPi ที่เกิดขึ้น ให้เป็นสารประกอบพลังงาน ATP ในสถานะที่มี adenosine 5'-phosphosulfate (APS) จากนั้นเอนไซม์ลูซิเฟอเรสจะใช้พลังงานจาก ATP เปลี่ยน ลูซิเฟอริน (luciferin) ให้เป็นออกซิลูซิเฟอริน (oxyluciferin) พร้อมกับปลดปล่อยแสงออกมา โดยปริมาณแสงที่เกิดขึ้นจะเป็นสัดส่วนกับปริมาณ ATP ที่เกิดขึ้น สัญญาณแสงจะถูกตรวจจับและบันทึกด้วยกล้อง CCD สัญญาณแสงที่เกิดขึ้นสามารถบ่งชี้ถึงสัดส่วนของจำนวนนิวคลีโอไทด์ที่ถูกนำไปต่อสายในการสังเคราะห์ดีเอ็นเอ โปรแกรมจะสร้างกราฟแสดงความเข้มของแสงที่อ่านได้ เรียกว่า Flowgram ของแต่ละหลุมใน PTP หลังจากนั้นเอนไซม์อะไพเรสจะทำหน้าที่ย่อยสลาย dNTP และ ATP ที่เหลืออยู่ในปฏิกิริยา เมื่อย่อยสลายสมบูรณ์แล้ว

dNTP ตัวใหม่จะถูกเติมลงไปเพื่อทำปฏิกิริยาในรอบต่อไป



รูปที่ 1 ขั้นตอนการวิเคราะห์ลำดับเบสด้วยเครื่อง Roche/454 sequencer (a) การเตรียมคลังดีเอ็นเอต้นแบบ โดยการตัดดีเอ็นเอจีโนมให้เป็นสายสั้น ๆ แล้วเชื่อมต่ออะแดปเตอร์ A และ B เข้าที่ปลายทั้งสองข้างของชิ้นดีเอ็นเอ (b) การเพิ่มปริมาณดีเอ็นเอด้วยเทคนิค emulsion PCR โดยการจับกันระหว่างดีเอ็นเอต้นแบบสายเดียวกับนิวคลีโอไทด์สายสั้น ๆ บนเม็ดบีด ในสัดส่วน 1:1 เพื่อให้หนึ่งเม็ดลูกบดจับกับดีเอ็นเอต้นแบบเพียงโมเลกุลเดียว แล้วเพิ่มปริมาณดีเอ็นเอในหยดน้ำมันที่ประกอบด้วยสารสำหรับการทำปฏิกิริยาพีซีอาร์ อยู่ใน (oil aqueous emulsion) (c) การวิเคราะห์ลำดับเบสไฟโรโดยการทำงานร่วมกันของเอนไซม์ DNA polymerase, ATP sulfurylase, luciferase และ apyrase ในแต่ละหลุมของ PicoTiterPlate<sup>TM</sup> [3]

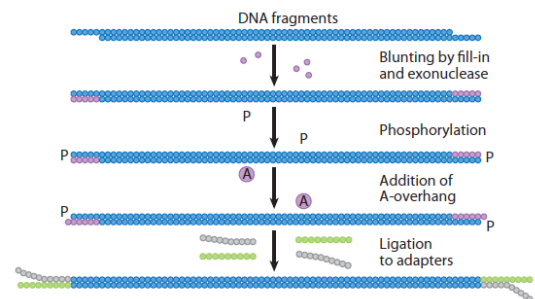
ข้อมูลลำดับเบสแต่ละเส้น (reads) ที่วิเคราะห์ได้จะมีความยาวของอยู่ในช่วง 100 ถึง 600 คู่เบส โดยในแต่ละครั้งของการวิเคราะห์ลำดับเบสจะได้จำนวนของเบสทั้งหมดอยู่ในช่วง 400 ถึง 600 ล้านเบส และความถูกต้องของเบสที่ได้มากกว่าร้อยละ 99 เนื่องจากแต่ละรอบของปฏิกิริยามีการเติม dNTP ลงไปเพียงชนิดเดียว ดังนั้นการวิเคราะห์ลำดับเบสไฟโรจึงมีข้อผิดพลาดจากการอ่านลำดับเบสต่ำ แต่อย่างไรก็ตาม ข้อจำกัดที่สำคัญของวิธีการนี้เกิดจากการอ่านลำดับเบสจากโมเลกุลต้นแบบที่เป็นลำดับเบสซ้ำหนึ่งเบส (homopolymer) เช่น AAAA ทำให้เกิดการปลดปล่อยสัญญาณแสงที่เข้มข้น ซึ่งยากต่อระบบในการตรวจจับและแปลงสัญญาณความเข้มเป็นจำนวนเบสที่ถูกต้องสำหรับลำดับเบสซ้ำหนึ่งเบส ที่มีขนาดยาวกว่า 8 หรือ 9 เบส ส่งผลให้การอ่านลำดับเบสผิดพลาดทำให้ลำดับเบสที่อ่านได้มีความยาวแตกต่างจากสายต้นแบบ

เครื่อง 454/Roche genome sequencer จะให้ข้อมูลที่อยู่ในรูปแบบของไฟล์ SFF (standard flowgram format) ซึ่งประกอบด้วยข้อมูลความเข้มแสง และข้อมูลลำดับเบสซึ่งแต่ละเบสจะกำกับด้วยคะแนนคุณภาพเฟด (Phred quality score) ซึ่งเป็นค่าที่แสดงระดับความถูกต้องของข้อมูลเบสที่วิเคราะห์ได้ โดยรูปแบบไฟล์ SFF สามารถอ่านได้ด้วยโปรแกรมของบริษัทที่ให้มาพร้อมกับเครื่องวิเคราะห์ลำดับเบส สำหรับโปรแกรมที่เป็นสาธารณะ เช่น Flower เป็นโปรแกรมที่ถูกพัฒนาขึ้นด้วยภาษา Haskell หรือ SFF Workbench ซึ่งเป็นโปรแกรมที่ถูกพัฒนาขึ้นโดย Heracle BioSoft S.R.L. ทั้งสองโปรแกรมสามารถอ่านไฟล์ .SFF และสามารถแปลงให้เป็นข้อมูลลำดับเบสที่อยู่ในรูปแบบของ FASTA หรือ FASTQ ได้ซึ่งเป็นรูปแบบของข้อมูลลำดับเบสที่ใช้ในการวิเคราะห์ในขั้นตอนต่อไป และเป็นรูปแบบมาตรฐานของข้อมูล

ลำดับเบสที่ได้จากเทคโนโลยีการวิเคราะห์ลำดับเบสยุคใหม่

## 2.2 Illumina Genome Analyzer

เทคโนโลยีการวิเคราะห์ลำดับเบสด้วยเครื่อง Illumina Genome Analyzer ถูกนำออกมาสู่การค้าเป็นครั้งแรกในปี ค.ศ. 2006 โดยบริษัท Solexa ถูกพัฒนาขึ้นโดย Shankar Balasubramanian และ David Klenerman ที่มหาวิทยาลัยเคมบริดจ์ และต่อมาปี ค.ศ. 2007 ถูกซื้อโดยบริษัท Illumina การวิเคราะห์ลำดับเบสด้วยเครื่อง Illumina Genome Analyzer ใช้หลักการของการสังเคราะห์สายดีเอ็นเอในการหาลำดับเบส ร่วมกับการใช้นิวคลีโอไทด์ที่มีการดัดแปลงให้มีคุณสมบัติหยุดการสังเคราะห์สายดีเอ็นเอและคืนสภาพกลับเพื่อให้ต่อสายดีเอ็นเอได้ (reversible terminator nucleotide) นิวคลีโอไทด์ดัดแปลงนี้ติดฉลากฟลูออเรสเซนต์

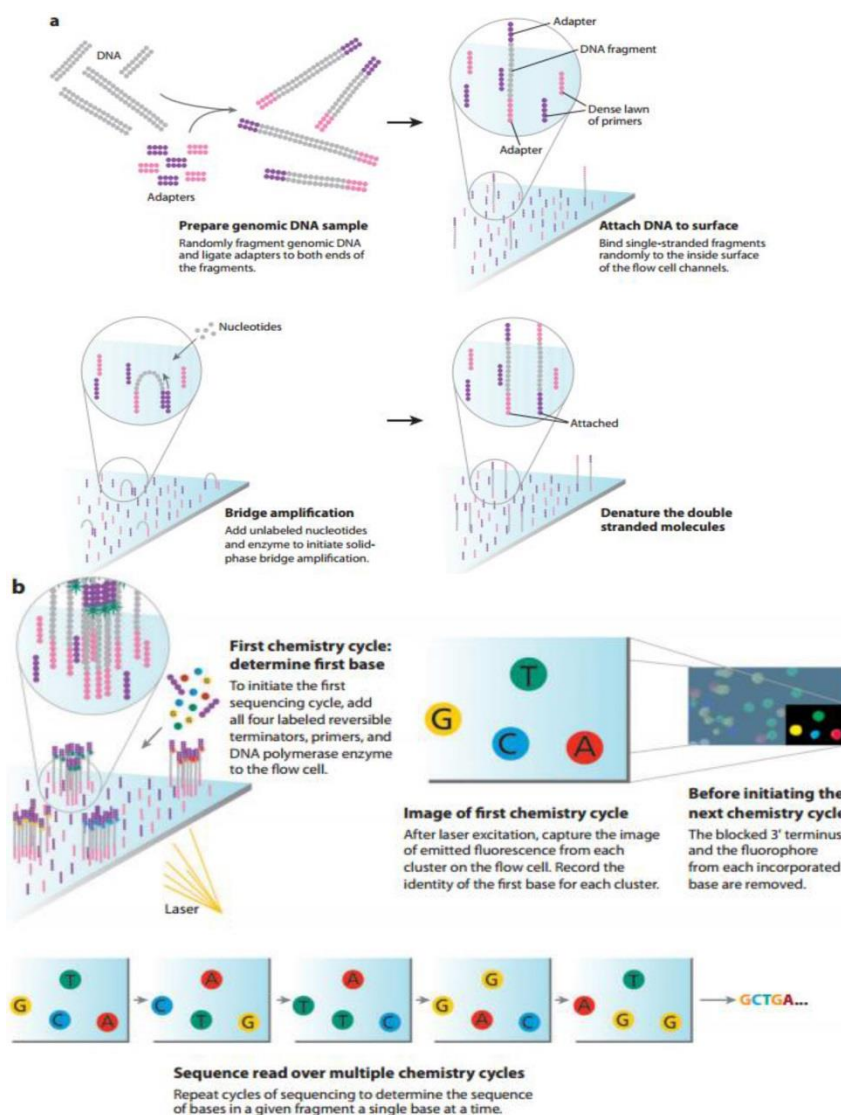


รูปที่ 2 การเตรียมดีเอ็นเอต้นแบบ (DNA library) ของ Illumina [6]

รูปแบบการวิเคราะห์ลำดับเบสจะเตรียมคลัง ดีเอ็นเอต้นแบบโดยการเพิ่มปริมาณดีเอ็นเอต้นแบบบนแผ่นกระจกสไลด์ (flow cell) โดยการเพิ่มจำนวนแบบสะพานโค้งเชื่อมต่อ (bridge amplification) ทำโดยตัดดีเอ็นเอเป็นชิ้นที่ขนาดเล็ก ซ่อมแซมส่วนปลายของชิ้นดีเอ็นเอให้เป็นปลายทู่ [5] แล้ว

เชื่อมต่อกับ อะแดปเตอร์ที่ปลายทั้งสองข้าง (รูปที่ 2) จากนั้นคัดเลือกขนาดดีเอ็นเอที่มีความยาวประมาณ 150 ถึง 200 คู่เบส นำดีเอ็นเอต้นแบบไปสร้างคลังดีเอ็นเอด้วยวิธี bridge amplification โดยแยกสายดีเอ็นเอให้เป็นดีเอ็นเอสายเดี่ยวแล้วนำไปตรึงบนแผ่นกระจกสไลด์ flow cell ซึ่งบนกระจกสไลด์จะเคลือบด้วยโอลิโก-นิวคลีโอไทด์ที่เป็นเบสคู่สมกับอะแดปเตอร์ ทำให้ชั้น ดีเอ็นเอต้นแบบติดบนแผ่น

กระจกสไลด์ แล้วเพิ่มจำนวนดีเอ็นเอด้วยเทคนิคพีซีอาร์ โดยดีเอ็นเอสายเดี่ยวหนึ่งโมเลกุลจะจับกับโอลิโกนิวคลีโอไทด์บนแผ่นกระจกสไลด์ซึ่งทำหน้าที่เป็นไพรเมอร์เป็นรูปสะพานโค้งเชื่อมต่อ (รูปที่ 3a) แล้วเพิ่มจำนวนดีเอ็นเอเป็นสายคู่ จากนั้นแยกดีเอ็นเอเป็นสายเดี่ยว เพื่อใช้เป็นคลังดีเอ็นเอต้นแบบสำหรับการวิเคราะห์ลำดับเบสต่อไป



รูปที่ 3 (a) การสร้างคลังดีเอ็นเอด้วยวิธี bridge amplification (b) การหาลำดับเบสโดยอาศัยหลักการการสังเคราะห์สายดีเอ็นเอร่วมกับการใช้ reversible dye terminators [7]

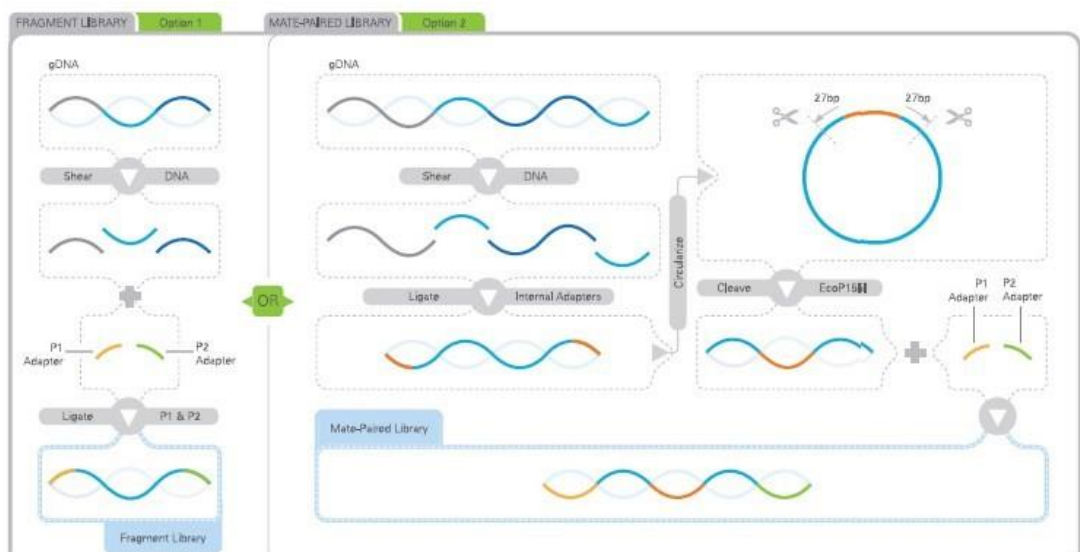
เมื่อได้คลังดีเอ็นเอต้นแบบสายเดี่ยวแล้ว การหาลำดับเบสของเครื่อง Illumina จะเติมไพรเมอร์ และนิวคลีโอไทด์ทั้งสี่ชนิดที่ติดฉลากด้วยสารเรืองแสง ด้วยสีที่ต่างกัน รวมถึงเติมเอนไซม์ดีเอ็นเอพอลิเมอร์ลงในปฏิกิริยา โดยนิวคลีโอไทด์ที่เติมจะมีคุณสมบัติหยุดการสังเคราะห์สายดีเอ็นเอ คือมีหมู่เคมีที่ป้องกันการต่อสายที่ปลาย 3' OH ไว้ ทำให้กระบวนการสังเคราะห์ดีเอ็นเอหยุดลง และนิวคลีโอไทด์ที่ไม่ได้เข้าคู่กับดีเอ็นเอต้นแบบที่เหลือจะถูกล้างออก แล้วบันทึกภาพการปลดปล่อยสารเรืองแสงของนิวคลีโอไทด์ที่เข้าคู่กับดีเอ็นเอต้นแบบ จากนั้นเกิดกระบวนการตัดหมู่เคมีที่ป้องกันการต่อสายที่ปลาย 3' OH และสารเรืองแสงออก ทำให้นิวคลีโอไทด์กลับมาเป็นนิวคลีโอไทด์ปกติ จึงเรียกว่า reversible terminator nucleotide หลังจากนั้นจะเติมนิวคลีโอไทด์ชุดใหม่ และเอนไซม์ดีเอ็นเอพอลิเมอร์เพื่อสังเคราะห์สายดีเอ็นเอในรอบต่อไป (รูปที่ 3b)

ความยาวของสายดีเอ็นเอที่ได้จากเครื่อง Illumina จะมีขนาดเท่ากันหมดโดยสามารถอ่านความ

ยาวได้ถึง 150 คู่เบส ในการหาลำดับเบสหนึ่งครั้งจะได้จำนวนของเบส 600 กิกะเบส และสามารถหาลำดับเบสบริเวณที่เป็นลำดับเบสซ้ำ (homopolymer) ได้อย่างแม่นยำ เครื่อง Illumina เหมาะกับการวิเคราะห์ลำดับเบสจีโนมเพื่อค้นหา SNPs (single nucleotide polymorphisms) เนื่องจากจำนวนสายดีเอ็นเอที่ได้มีจำนวนมากและเป็นสายดีเอ็นเอขนาดสั้น ทำให้มีจำนวนซ้ำการวิเคราะห์เบสตำแหน่งเดิม (coverage) สูง ดังนั้นการค้นหา SNPs จึงทำได้อย่างมีประสิทธิภาพ ข้อมูลที่ได้จากเครื่องจะเป็นรูปแบบมาตรฐาน คือ FASTQ หรือ BAM [8]

### 2.3 Applied Biosystems SOLiD™ Sequencer

เทคโนโลยีการหาลำดับเบสของ Applied Biosystems SOLiD™ Sequencer หรือเรียกสั้น ๆ ว่า SOLiD ถูกนำออกสู่การค้าในปี ค.ศ. 2006 โดยบริษัท Applied Biosystems ใช้หลักการหาลำดับเบสโดยการต่อเชื่อมสายดีเอ็นเอด้วยปฏิกิริยาไลเกชัน (sequencing by ligation) โดยใช้โพรบสองเบส ที่ติด



รูปที่ 4 การสร้างคลังดีเอ็นเอแบบชิ้นดีเอ็นเอ (fragment library) และการสร้างคลังดีเอ็นเอแบบจับคู่ปลายชิ้นดีเอ็นเอ (mate-paired library) [9]

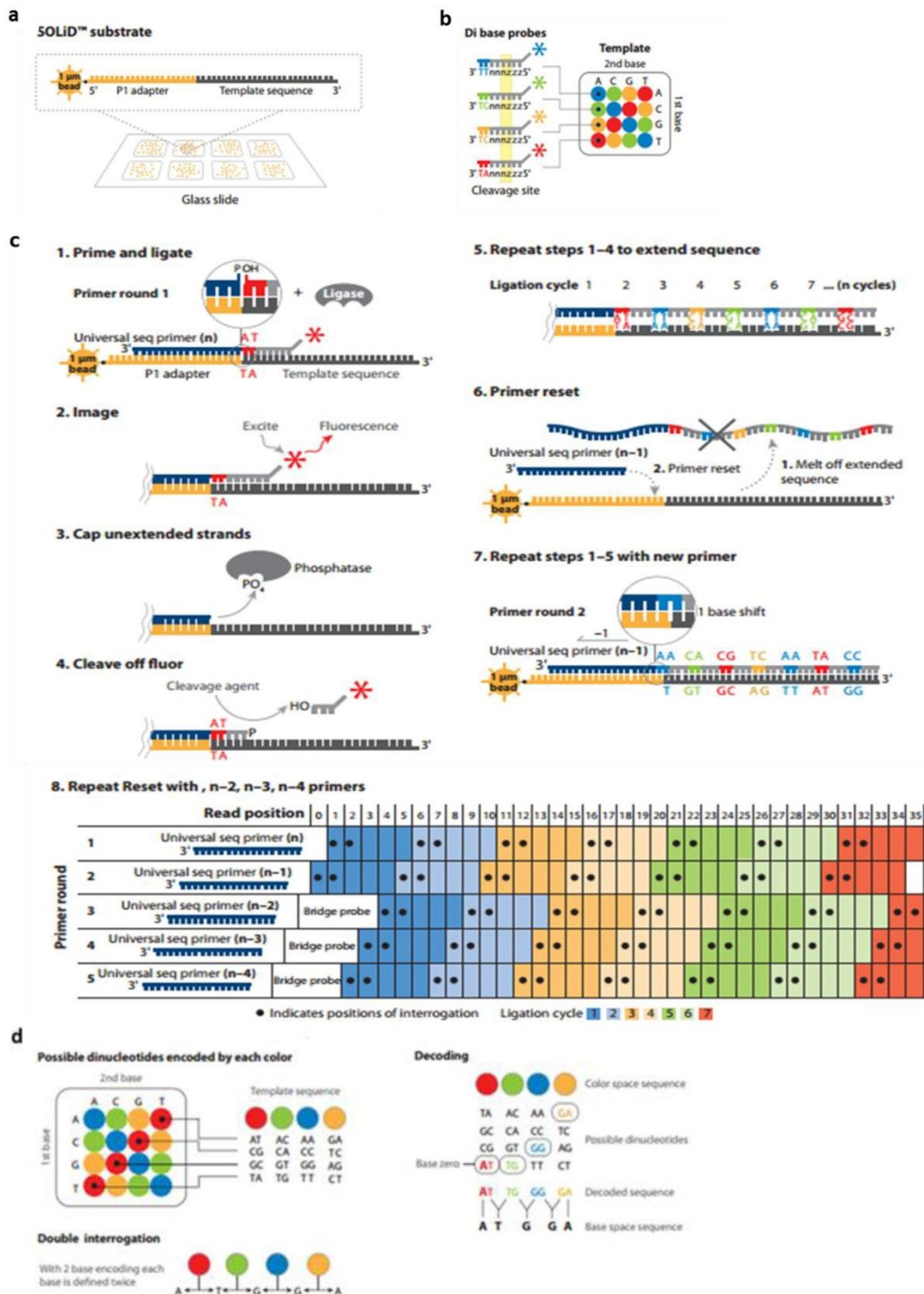
ฉลากด้วยสารเรืองแสง (di-base probe) ด้วยหลักการของโพรบสองเบสสามารถเข้าคู่กับดีเอ็นเอต้นแบบได้อย่างจำเพาะ การสร้างคลังดีเอ็นเอต้นแบบสำหรับการวิเคราะห์ลำดับเบสของ SOLiD ใช้การเพิ่มปริมาณดีเอ็นเอด้วยอิมัลชันพีซีอาร์ บนเม็ดลูกปัดทำนองคล้ายกับ 454/Roche

ในการเตรียมคลังดีเอ็นเอต้นแบบ ทำโดยตัดดีเอ็นเอให้มีขนาดยาวประมาณ 200 คู่เบส การเตรียมดีเอ็นเอต้นแบบทำได้สองวิธี (รูปที่ 4) ได้แก่ คลังดีเอ็นเอแบบชิ้นดีเอ็นเอ (fragment library) เป็นการเชื่อมต่อกับอะแดปเตอร์ที่ปลายทั้งสองด้านของชิ้นดีเอ็นเอทำนองเดียวกับการเตรียมคลังดีเอ็นเอทั่วไป และอีกแบบหนึ่งคือคลังดีเอ็นเอแบบจับคู่ปลายชิ้นดีเอ็นเอ (mate-paired library) [9] โดยเชื่อมต่อกับอะแดปเตอร์ที่ปลายทั้งสองด้านของชิ้นดีเอ็นเอ จากนั้นนำปลายของอะแดปเตอร์มาเชื่อมต่อกันได้เป็นโมเลกุลวงแหวน ตัดโมเลกุลวงแหวนสองตำแหน่งได้ ชิ้นดีเอ็นเอที่มีปลายทั้งสองด้านเป็นดีเอ็นเอที่ต้องการวิเคราะห์ลำดับเบส จากนั้นต่ออะแดปเตอร์ที่ปลายหัวท้ายทำนองเดียวกับคลังดีเอ็นเอแบบชิ้นดีเอ็นเอ ดีเอ็นเอสายคู่จะถูกทำให้เป็นดีเอ็นเอสายเดี่ยว นำดีเอ็นเอสายเดี่ยวที่เตรียมได้ตรงติดเข้ากับเม็ดลูกปัดที่เคลือบด้วยโอลิโกนิวคลีโอไทด์ที่เป็นคู่สมกับอะแดปเตอร์ แล้วเพิ่มจำนวนโมเลกุลของดีเอ็นเอต้นแบบด้วยการทำอิมัลชันพีซีอาร์ แยกเม็ดลูกปัดที่ไม่มีการเพิ่มจำนวนดีเอ็นเอออก แล้วตรึงโมเลกุลดีเอ็นเอที่ติดกับเม็ดลูกปัดบนแผ่นกระจกสไลด์ (รูปที่ 5a) เพื่อใช้ในการหาลำดับเบสต่อไป

การหาลำดับเบสเริ่มโดยเติมไพรเมอร์ซึ่งเป็นคู่สมกับอะแดปเตอร์ที่เชื่อมกับดีเอ็นเอต้นแบบ SOLiD ใช้หลักการหาลำดับเบสที่แตกต่างจากระบบการวิเคราะห์ลำดับเบสอื่น คือไม่ใช่เอนไซม์ดีเอ็นเอพอลิเมอเรส แต่ใช้การเชื่อมต่อเชื่อมสายดีเอ็นเอด้วย

ปฏิกิริยาไลเกชัน โดยใช้โพรบสองเบส 16 แบบ โพรบสองเบสติดฉลากด้วยสารเรืองแสงแตกต่างกัน 4 สี หนึ่งสัระบุโพรบสองเบสสองแบบ โพรบสองเบสนี้มีส่วนต่อเพิ่มที่ปลาย 3' อีกหกเบสรวมเป็นความยาวแปดเบส (รูปที่ 5b) โดยเบสที่หนึ่งและสอง เป็นเบสจำเพาะที่สามารถเข้าคู่กับดีเอ็นเอต้นแบบได้ เมื่อเติมไพรเมอร์ลงในปฏิกิริยาแล้วเติมโพรบสองเบส ทั้ง 16 แบบ โพรบสองเบสที่สามารถเข้าคู่กับดีเอ็นเอต้นแบบจะสามารถสร้างพันธะไฮโดรเจนกับเบสของดีเอ็นเอต้นแบบได้ จากนั้นเอนไซม์ไลเกสจะเชื่อมต่อโพรบสองเบสกับไพรเมอร์ โพรบที่เหลือที่ไม่สามารถเข้าคู่กับดีเอ็นเอต้นแบบได้อย่างจำเพาะจะถูกล้างออก แล้วเกิดการบันทึกภาพสารเรืองแสงที่ติดกับโพรบที่เชื่อมต่อกับไพรเมอร์ ต่อมาจะเกิดกระบวนการตัดโพรบที่ตำแหน่งระหว่างเบสที่ห้าและหกออก แล้วเข้าสู่การเชื่อมรอบต่อไป โดยการเติมโพรบสองเบสรอบใหม่เข้าไป บันทึกภาพ และการกำจัดโพรบที่เหลือ กระบวนการจะซ้ำไปเรื่อยๆจนได้ความยาวสายดีเอ็นเอที่ต้องการ จากนั้นทำการเริ่มต้นใหม่ (reset) โดยใช้ไพรเมอร์เดิมที่ลดจำนวนเบสที่ปลาย 3' ไปหนึ่งตัว (n-1) ให้เข้าจับกับอะแดปเตอร์ ซึ่งจำนวนรอบของการเชื่อม บันทึกภาพ และการกำจัดโพรบที่เหลือ เกิดขึ้นซ้ำ ๆ หลังจากนั้นจะเริ่มต้นใหม่โดยใช้ไพรเมอร์เป็น n-2, n-3 และ n-4 (รูปที่ 5c) ทำให้ได้ข้อมูลที่มีการเชื่อมซ้อนกันสามารถหาลำดับเบสจากสี่ที่บันทึกได้

สำหรับการอ่านลำดับเบสจากสี่จะต้องทราบนิวคลีโอไทด์แรกของอะแดปเตอร์ก่อน แล้วจากนั้นจะแปลลำดับสี่เป็นลำดับเบสจากตารางสี่ที่บอกถึงความเป็นไปได้ของโพรบสองเบส ทั้ง 16 แบบ ที่ให้สีต่างกัน (รูปที่ 5d) เช่น ถ้านิวคลีโอไทด์ตัวแรกของลำดับเบสอะแดปเตอร์เป็น A และสี่ที่บันทึกได้สี่แรกคือสีแดง เบสต่อไปเมื่อเทียบในตารางสี่ของโพรบสองเบส 16 แบบ จะเป็นเบส T ต่อมาสี่ที่บันทึกได้เป็นสี่ที่



รูปที่ 5 (a) การเตรียมคลั่งดีเอ็นเอต้นแบบสำหรับการหาลำดับเบส (b) di-base probe ซึ่งมีทั้งหมด 16 แบบ ที่สามารถแปลเป็นสี่ของสารเรืองแสงฟลูออเรสเซนซ์ได้สี่สี (c) การหาลำดับเบสด้วยวิธีการต่อเชื่อมสายดีเอ็นเอด้วยปฏิกิริยาไลเกชัน (d) การแปลรหัสสี่สีที่บันทึกได้เป็นลำดับเบส [9]

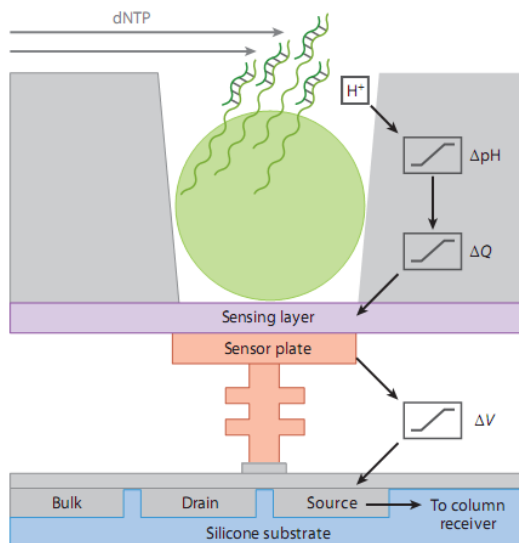
สองคือสี่เขียว ดังนั้นเบสลำดับต่อไปก็จะเป็นเบส G และถ้าต่อมาสี่ที่บันทึกได้เป็นสีฟ้า เบสต่อไปจะเป็นเบส G เป็นต้น (รูปที่ 5d)

ความยาวสายดีเอ็นเอที่ได้จากเครื่อง SOLiD ยาวประมาณ 35 ถึง 75 คู่เบส และการหาลำดับเบสหนึ่งครั้งจะได้จำนวนเบสประมาณ 100 ถึง 250 กิกะเบส เครื่อง SOLiD เหมาะสำหรับการค้นหา SNPs เนื่องจากสามารถแยกแยะความผิดพลาดของการหาลำดับเบสกับ SNPs ได้ โดยหากมีความผิดพลาดในของการหาลำดับเบสเกิดขึ้นเมื่อเปรียบเทียบกับลำดับเบสอ้างอิงแล้วสี่ที่บันทึกได้จะแตกต่างเพียงหนึ่ง สี แต่หากเป็น SNPs จะมีความแตกต่างของสีสองสี [8]

## 2.4 Ion Torrent/ Life technology

เทคโนโลยีการวิเคราะห์ลำดับเบสด้วยเครื่อง Ion Personal Genome Machine (PGM) หรือที่รู้จักกันในชื่อ Ion torrent sequencing, pH-mediated sequencing, silicon sequencing หรือ semiconductor sequencing เป็นรูปแบบการวิเคราะห์ลำดับเบสยุคใหม่ที่เกิดขึ้นในช่วงปลายปี ค.ศ. 2010 เป็นวิธีการวิเคราะห์ลำดับเบสที่มีความคล้ายคลึงกับเทคนิค การวิเคราะห์ลำดับเบสไฟโรดิงที่กล่าวไปแล้วข้างต้น เทคนิคนี้เป็นการตรวจวัดการเปลี่ยนแปลงของค่า pH ที่เกิดจากการปลดปล่อยโปรตอน ( $H^+$ ) ในระหว่างการสังเคราะห์สายดีเอ็นเอที่เกิดจากการทำงานของเอนไซม์ดีเอ็นเอพอลิเมอเรส ดังนั้นจึงเป็นวิธีการวิเคราะห์ลำดับเบสโดยการสังเคราะห์สายดีเอ็นเอ เทคนิคนี้ถูกสร้างโดยใช้หลักการเดียวกันกับเทคโนโลยีของสารกึ่งตัวนำ (semiconductor technology) ช่วยให้สามารถแปลผลข้อมูลทางพันธุกรรมให้เป็นข้อมูลดิจิทัล (ลำดับดีเอ็นเอ) ได้โดยตรง มีลักษณะเป็นแผ่นชิป (semiconductor chip) ขนาดเล็กที่ประกอบด้วยกรวยตัวของหลุม

เล็ก ๆ (micro wells) ประมาณ 400 ช่อง ต่อความกว้างเท่ากับเส้นผมของคน วางตัวอยู่บนชั้นที่มีความไวต่อไอออนสูง (ion-sensitive layer) และตัวตรวจวัดไอออน (ion sensor) [10,11] (รูปที่ 6)



รูปที่ 6 ภาพตัดขวางของหลุม (microwell) บนแผ่นชิป (semiconductor chip) ภายในหลุมบรรจุ Ion Sphere™ particles ที่มีดีเอ็นเอต้นแบบจับอยู่บนพื้นผิวโมเลกุล เมื่อมีการเข้าสู่ของนิวคลีโอไทด์ จะเกิดการปลดปล่อย  $H^+$  และทำให้เกิดการเปลี่ยนแปลงของค่าพีเอชในแต่ละหลุม ซึ่งถูกตรวจวัดได้ด้วยเซ็นเซอร์ จากนั้นสัญญาณทางเคมีที่เกิดขึ้นจะเปลี่ยนเป็นสัญญาณดิจิทัลซึ่งก็คือลำดับดีเอ็นเอ [6]

การเตรียมดีเอ็นเอต้นแบบ และการเพิ่มปริมาณดีเอ็นเอสำหรับการวิเคราะห์ลำดับเบสใช้หลักการเดียวกับ 454 pyrosequencing โดยตัดดีเอ็นเอจีโนมแบบสุ่ม ให้เป็นสายสั้นๆ แล้วเชื่อมต่ออะแดปเตอร์ที่ปลายทั้งสองข้างของชิ้นดีเอ็นเอ จากนั้นนำโมเลกุลดีเอ็นเอเข้าจับกับเม็ดลูกปัด (Ion Sphere™

particles) เพื่อเพิ่มปริมาณดีเอ็นเอต้นแบบด้วยเทคนิคอิมัลชันพีซีอาร์ ขั้นตอนการวิเคราะห์ลำดับเบสเริ่มจากการนำเม็ดยูทูปที่มีดีเอ็นเอต้นแบบติดอยู่ใส่ลงในหลุมขนาดเล็ก (microwell) บนแผ่นชิป (semiconductor chip) แล้วนำเข้าเครื่องวิเคราะห์ลำดับเบส ภายในเครื่องจะมีการเติมสารสำหรับการสังเคราะห์สายดีเอ็นเอแบบอัตโนมัติ โดยแต่ละรอบของปฏิกิริยาจะเติม dNTP เพียงชนิดเดียว เมื่อเกิดการเข้าคู่ของเบสคู่สมกับดีเอ็นเอต้นแบบ การทำงานของเอนไซม์ดีเอ็นเอพอลิเมอเรสจะทำให้เกิดการปลดปล่อย H<sup>+</sup> ซึ่งทำให้เกิดการเปลี่ยนแปลงของค่า pH ที่สามารถตรวจวัดได้ด้วยตัวตรวจวัดไอออน และแปลผลออกมาเป็นข้อมูลลำดับเบส

ข้อมูลลำดับเบสที่วิเคราะห์ได้จะอยู่ในรูปแบบของไฟล์ SFF และ FASTQ ซึ่งง่ายต่อการนำไปวิเคราะห์ในขั้นตอนต่อไป มีความยาวของสายดีเอ็นเออยู่ในช่วง 200 ถึง 400 คู่เบส โดยในแต่ละครั้งของการวิเคราะห์ลำดับเบสจะได้จำนวนของเบสทั้งหมดอยู่ในช่วง 100 ถึง 10,000 ล้านเบส ขึ้นอยู่กับชนิดของแผ่นชิปที่ใช้ ได้แก่ Ion 314™, Ion 316™, Ion 318™, และ Ion PI™ เทคโนโลยีการวิเคราะห์ลำดับเบสนี้แตกต่างจากเทคโนโลยีอื่น ๆ เนื่องจากไม่ต้องมีการดัดแปลงนิวคลีโอไทด์ที่ใช้ในการสังเคราะห์ดีเอ็นเอ ไม่ต้องอาศัยการทำงานของเอนไซม์หลายชนิด และไม่ต้องอาศัยเครื่องมือสำหรับการตรวจจับสัญญาณแสงที่เกิดขึ้น จึงทำให้สามารถลดต้นทุนในการวิเคราะห์ลำดับเบส และทำให้การวิเคราะห์ลำดับเบสมีความรวดเร็วยิ่งขึ้น แต่ข้อจำกัดที่สำคัญของเทคนิคนี้คือ อาจเกิดข้อผิดพลาดการวิเคราะห์ลำดับเบสซ้ำชนิดหนึ่งเบส (homopolymer) เช่นเดียวกับเทคนิคการวิเคราะห์ลำดับเบสไพโร เนื่องจากการต่อสายของนิวคลีโอไทด์มากกว่าหนึ่งตัว ในหนึ่งรอบของการทำปฏิกิริยา ทำให้เกิดการปลดปล่อย H<sup>+</sup> ที่มากขึ้น ส่งผล

ให้การเปลี่ยนแปลงค่า pH มากขึ้นตามไปด้วย โดยการเปลี่ยนแปลงค่า pH ที่สูงมากเกินไปทำให้ยากต่อการแยกความแตกต่างระหว่างเบสซ้ำที่มีจำนวนเบสมากกว่าแปดตัวขึ้นไป

### 3. เทคโนโลยีการวิเคราะห์ลำดับเบสยุคที่สาม

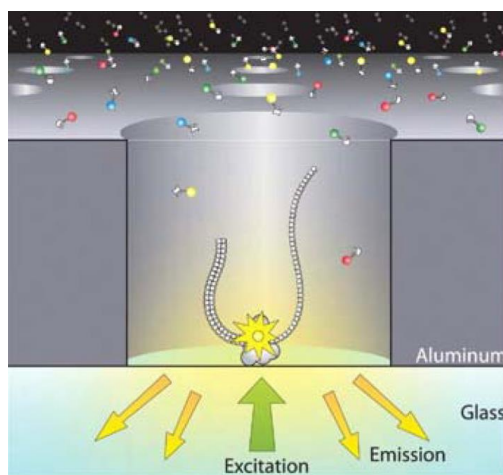
TGSTs มีทั้งที่สามารถผลิตและขายในเชิงการค้าได้ แต่บางเทคนิคก็อยู่ระหว่างขั้นตอนการพัฒนา TGSTs ที่สามารถผลิตในเชิงการค้าได้มีสองแบบ แบบที่หนึ่งเป็นการวิเคราะห์ลำดับเบสโดยใช้การสังเคราะห์ดีเอ็นเอ (sequencing by synthesis) จากดีเอ็นเอธรรมชาติโมเลกุลเดี่ยว กระบวนการสังเคราะห์ดีเอ็นเอโดยเอนไซม์ดีเอ็นเอพอลิเมอเรสจะถูกบันทึกตามเวลาจริง แบบที่สองเป็นเทคโนโลยีการวิเคราะห์ลำดับเบส โดยนำสายดีเอ็นเอผ่านรูขนาดเส้นผ่านศูนย์กลางระดับนาโนเมตร (nanopore sequencing) นิวคลีโอไทด์แต่ละตัวในองค์ประกอบของสายดีเอ็นเอเมื่อเคลื่อนผ่านรูที่มีขนาดเส้นผ่านศูนย์กลางระดับนาโนเมตรจะถูกบันทึก

เทคโนโลยีการวิเคราะห์ลำดับเบสยุคที่สามแบบที่ใช้วิธีการสังเคราะห์ดีเอ็นเอพัฒนาโดย Pacific Biosciences ([www.pacificbiosciences.com](http://www.pacificbiosciences.com)) มีชื่อว่า PacBio® RS (PacBio® RS High Resolution Genetic Analyzer) เครื่องวิเคราะห์ลำดับเบสรุ่นนี้ใช้เทคโนโลยี SMRT (Single Molecule, Real-Time [12]) เป็นการวิเคราะห์ดีเอ็นเอธรรมชาติโมเลกุลเดี่ยวแบบตามเวลาจริงบนแผ่นเซลล์ SMRT ที่ประกอบด้วย ZMW (zero mode waveguides) [13] หลักการทำงานของเทคโนโลยีนี้คือการตรวจจับการเข้าต่อสายของนิวคลีโอไทด์ที่ติดฉลากฟลูออเรสเซนต์สีต่างกันสี่สี ขณะเอนไซม์ดีเอ็นเอพอลิเมอเรสกำลังสังเคราะห์ดีเอ็นเอบนแผ่นเซลล์ SMRT ที่มีหลุมขนาดเล็กของ ZMW จำนวน 150,000 ZMW ต่อหนึ่งเซลล์ ZMW เป็นหลุม

ขนาดเล็กมีเส้นผ่านศูนย์กลางประมาณ 10 นาโนเมตร สร้างบนแผ่นฟิล์มโลหะบนผิวหน้าของแผ่นซิลิกาหรือแผ่นแก้ว (รูปที่ 7) ด้วยขนาดของหลุมที่เล็กทำให้การเรืองแสงฟลูออเรสเซนซ์ถูกจำกัดอยู่บริเวณกันหลุมจะไม่ไปรบกวนบริเวณด้านบนของหลุม ในเทคโนโลยี SMRT นี้ เอนไซม์ดีเอ็นเอพอลิเมอเรสหนึ่งโมเลกุลจะถูตรึงอยู่ที่กันหลุมบนแผ่นแก้วโดยใช้ไบโอติน/สเตรปตาวิดิน (biotin/streptavidin) ในระหว่างรอบการวิเคราะห์ลำดับเบส ดีเอ็นเอต้นแบบจะถูกใส่ลงไปในหลุมที่มีเอนไซม์ดีเอ็นเอพอลิเมอเรส จากนั้นนิวคลีโอไทด์ที่ติดฉลากฟลูออเรสเซนซ์จะถูกนำไปต่อโดยเอนไซม์ดีเอ็นเอพอลิเมอเรส กระบวนการนี้จะถูกตรวจจับตามเวลาจริงโดยระบบการจับภาพของ ZMW นิวคลีโอไทด์ในเทคโนโลยี SMRT นี้ติดฉลากฟลูออเรสเซนซ์ที่หมู่ฟอสเฟต หน่วยนิวคลีโอไทด์เพียงหนึ่งหน่วยจะสามารถแพร่ผ่าน ZMW และถูกนำไปต่อสายดีเอ็นเอในหนึ่งหลุมของ ZMW ส่วนของฟลูออเรสเซนซ์จะถูกตัดออกจากหมู่ฟอสเฟตตามธรรมชาติของการสังเคราะห์ดีเอ็นเอโดยเอนไซม์ดีเอ็นเอพอลิเมอเรส ระบบของ ZMW จะทำการตรวจจับแสงฟลูออเรสเซนซ์ในรอบต่อไปเมื่อเอนไซม์ดีเอ็นเอพอลิเมอเรสมีที่ว่าง นิวคลีโอไทด์ตัวต่อไปจะถูกเลือกโดยเอนไซม์ดีเอ็นเอพอลิเมอเรส และเข้าสู่รอบการสังเคราะห์ดีเอ็นเอรอบใหม่ ระบบสามารถตรวจจับได้ 75,000 ZMW ในเวลาเดียวกัน

PacBio® RS ใช้เอนไซม์ดีเอ็นเอพอลิเมอเรสที่มีประสิทธิภาพสูงในการสังเคราะห์ดีเอ็นเอคือสามารถสังเคราะห์ดีเอ็นเอได้ยาวประมาณ 2,000 นิวคลีโอไทด์ ปัจจัยสำคัญที่จำกัดการทำงานของเอนไซม์คือการที่เอนไซม์ถูกทำลายเนื่องจากแสงเลเซอร์ที่ใช้ในการจับภาพ ข้อจำกัดที่สำคัญอีกประการหนึ่งของ SMRT คือประสิทธิภาพในการจับภาพของระบบ ZMW ต่อครั้งหนึ่งของการวิเคราะห์ ทำให้ระบบสามารถผลิตข้อมูล

ได้ประมาณ 40-90 เมกาเบส ซึ่งเป็นปัจจัยที่ต้องพัฒนาต่อไป ข้อเสียที่สำคัญอีกประการหนึ่งของระบบ SMRT คือการเกิดข้อผิดพลาดของการวิเคราะห์ในอัตราสูง 10 ถึง 15 เปอร์เซ็นต์ เป็นอัตราผิดพลาดในอัตราสูงที่สุดของการวิเคราะห์ลำดับเบสยุคใหม่ ทั้งนี้เนื่องจากประสิทธิภาพของการวิเคราะห์การต่อสายของโมเลกุลเดี่ยวที่เกิดขึ้นอย่างรวดเร็วในแบบเวลาจริงยังมีประสิทธิภาพไม่ดีมากนัก ซึ่งต้องมีการพัฒนาต่อไป แต่อย่างไรก็ตามในระบบนี้ใช้วิธีการเตรียมดีเอ็นเอต้นแบบที่มีลักษณะเป็นวงแหวน ดังนั้นการวิเคราะห์ลำดับเบสจึงสามารถทำได้แบบวนซ้ำ 10 ถึง 15 รอบของโมเลกุลเดิม จึงทำให้ลดอัตราความคลาดเคลื่อนได้ แต่ก็ทำให้ได้ข้อมูลลำดับเบสน้อยลงและได้ความยาวของสายดีเอ็นเอที่วิเคราะห์มีความยาวจำกัด



รูปที่ 7 เทคโนโลยีการวิเคราะห์ลำดับเบส PacBio® RS ดีเอ็นเอธรรมชาติต้นแบบโมเลกุลเดี่ยวถูกวิเคราะห์ลำดับเบสแบบเวลาจริง (real time) โดยการสังเคราะห์ดีเอ็นเอด้วยเอนไซม์ดีเอ็นเอพอลิเมอเรสที่ถูกตรึงที่กันหลุมของ ZMW นิวคลีโอไทด์ติดฉลากฟลูออเรสเซนซ์ที่ถูกนำไปต่อสายดีเอ็นเอจะปล่อยแสงฟลูออเรสเซนซ์ [13]

ลักษณะที่น่าสนใจประการหนึ่งของ SMRT คือระบบสามารถสังเกตและตรวจจับไคเนติก (kinetic) ของกิจกรรมของเอนไซม์ดีเอ็นเอพอลิเมอร์สขณะต่อสายดีเอ็นเอได้ ดังนั้นถ้ามีการเปลี่ยนแปลงไคเนติกของการต่อสายดีเอ็นเอเช่นดีเอ็นเอต้นแบบที่มีการเติมหมู่เมทิล (methylation) ระบบจะสามารถตรวจสอบได้ SMRT จึงสามารถนำมาใช้ศึกษาการเติมหมู่เมทิลของดีเอ็นเอได้ [14] นอกจากนี้ระบบยังมีความยืดหยุ่นของการประยุกต์ใช้งาน เช่นศึกษาการแปลรหัสเอ็มอาร์-เอ็นเอเป็นโปรตีนโดยไรโบโซมแบบเวลาจริง [15] สังเกตการทำงานของเอนไซม์รีเวิร์สทรานสคริปเตส (reverse transcriptase) แบบเวลาจริง จึงสามารถหาลำดับเบสของอาร์เอ็นเอโดยตรงได้

เทคโนโลยีการวิเคราะห์ลำดับเบสยุคที่สามแบบที่สอง เป็นเทคโนโลยีการวิเคราะห์ลำดับเบสโดยนำสายดีเอ็นเอผ่านรูขนาดเส้นผ่านศูนย์กลางระดับนาโนเมตรหรือนาโนพอร์ (nanopore) เทคโนโลยีนี้ได้เริ่มจากงานวิจัยของ Deamer และคณะ ณ มหาวิทยาลัยแห่งแคลิฟอร์เนีย [16] ต่อมา Oxford Nanopore Technologies® ได้เป็นผู้นำการพัฒนาการวิเคราะห์ลำดับเบสเทคโนโลยีนี้ โดยใช้หลักการผ่านกระแสไฟฟ้าข้ามรูขนาดเล็กระดับนาโนเมตร สนามไฟฟ้าจะทำให้เกิดแรงขับเคลื่อนโมเลกุลกรดนิวคลีอิกโครงสร้างสายเดี่ยวผ่านรูระดับนาโนเมตร การผ่านรูของนิวคลีโอไทด์แต่ละตัวจะถูกตรวจจับโดยการเปลี่ยนแปลงของกระแสไฟระหว่างเคลื่อนผ่านรูของนิวคลีโอไทด์เนื่องจากเทคโนโลยีนี้ใช้ดีเอ็นเอโมเลกุลเดี่ยวที่ไม่ผ่านกระบวนการเปลี่ยนแปลงแต่อย่างใด จึงเป็นเทคโนโลยีที่มีศักยภาพในการทำงานได้รวดเร็ว ใช้ปริมาณดีเอ็นเอน้อยมาก สามารถใช้กับการวิเคราะห์ลำดับเบสอาร์เอ็นเอโดยตรง หรือนิวคลีโอไทด์ที่ผ่านกระบวนการเปลี่ยนแปลง และกระบวนการวิเคราะห์ใช้เวลารวดเร็ว นอกจากนี้ข้อดีของการวิเคราะห์ลำดับเบสโดย

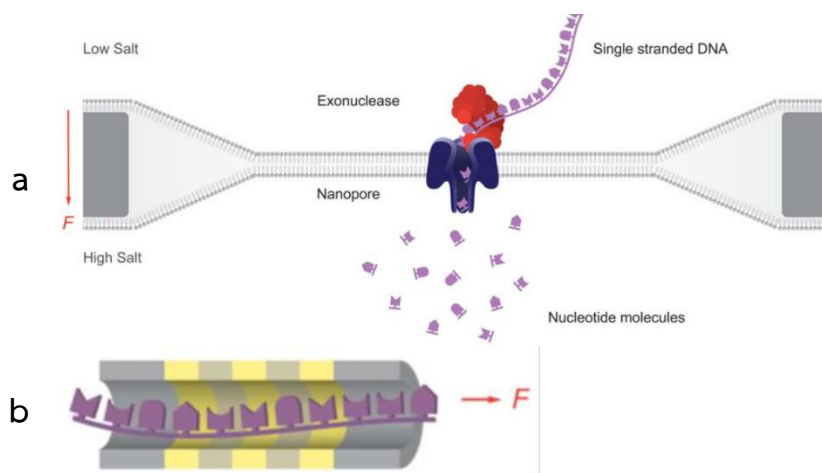
เทคโนโลยีนาโนพอร์ คือการไม่ต้องใช้กระบวนการสังเคราะห์ดีเอ็นเอเหมือนเทคโนโลยีอื่น และนาโนพอร์สามารถสามารถใช้ได้หลายครั้งจึงทำให้เทคโนโลยีนี้มีค่าใช้จ่ายน้อย

นาโนพอร์สามารถสร้างจากวัสดุทางชีวโมเลกุล เช่น เมมเบรนโครงสร้างเฮปตาเมอร์แอลฟาฮีโมไลซิน ( $\alpha$ -hemolysin, a transmembrane heptamer) ของแบคทีเรีย *Staphylococcus aureus* [17] ซึ่งเป็นระบบของ Oxford Nanopore Technologies® หรือวัสดุที่เป็นของแข็งเช่นกราฟีน (graphene) [18] และซิลิกอนไนเตรด (silicon nitride) ซึ่งในปัจจุบันได้รับความนิยมมากกว่าเนื่องจากความสามารถในการควบคุมโครงสร้างรู ทั้งความกว้างและความหนา และความสามารถในการผลิตในเชิงพาณิชย์

การควบคุมความเร็วในการผ่านรูของกรดนิวคลีอิกสามารถควบคุมได้หลายวิธี เช่น การเปลี่ยนแปลงความหนืดและความเข้มข้นของเกลือ [19] การเปลี่ยนแปลงทางเคมีของรูและเอนไซม์ดีเอ็นเอพอลิเมอร์ส [18] หรือเอนไซม์เอกโซนิวคลีเอส (exonuclease; รูปที่ 8a) [21] เพื่อช่วยในการเคลื่อนโมเลกุลผ่านรู การตรวจจับนิวคลีโอไทด์ที่เคลื่อนผ่านรูระบบส่วนมากจะใช้สัญญาณกระแสไฟฟ้าเนื่องจากมีต้นทุนต่ำ ผลิตง่าย แต่ก็มีผู้เสนอการตรวจจับนิวคลีโอไทด์ที่เคลื่อนผ่านรูด้วยระบบการจับภาพ (optical detection) [22] เทคโนโลยีหลังนี้พัฒนาโดย NobleGen ต้องมีการเปลี่ยนแปลงกรดนิวคลีอิกให้มีการติดฉลากฟลูออเรสเซนต์ที่เป็นรหัสซึ่งสามารถตรวจจับภาพได้เมื่อรหัสฟลูออเรสเซนต์เคลื่อนผ่านรู นอกจากนี้นักวิจัยจาก IBM ได้ทดลองสร้างดีเอ็นเอทรานซิสเตอร์ (DNA transistor; รูปที่ 8b) ซึ่งเป็นนาโนพอร์ชนิดวัสดุที่เป็นแผ่นชั้นของแข็งซ้อนกัน เบสแต่ละตัวบนดีเอ็นเอสายเดี่ยวที่เคลื่อนผ่านนาโนพอร์สามารถตรวจจับด้วยสัญญาณกระแสไฟฟ้าที่เกิดการ

เปลี่ยนแปลง ข้อดีของเทคโนโลยีนี้คือสามารถอ่านลำดับเบสโมเลกุลดีเอ็นเอสายยาวได้เร็ว และมีความคุ้มค่าเนื่องจากไม่มีกระบวนการติดฉลากดีเอ็นเอ และไม่ต้องมีระบบการตรวจจับภาพ โดยทางทฤษฎีสามารถอ่านลำดับเบสสายยาว 500 เมกะเบสในเวลาหนึ่งวินาทีต่อหนึ่งทรานซิสเตอร์ อย่างไรก็ตามข้อมูลมีการเผยแพร่สู่สาธารณะค่อนข้างจำกัดการพิสูจน์

แนวคิดนี้จึงยังไม่เป็นที่ยอมรับ ขณะที่เทคโนโลยีของ Oxford Nanopore Technologies<sup>®</sup> มีการพิสูจน์และสามารถผลิตในเชิงการค้าได้แล้ว โดยระบบจะถูกออกแบบให้สามารถใช้กับขนาดที่แตกต่างกันของงานวิจัยตั้งแต่โครงการวิเคราะห์ลำดับเบสจีโนมขนาดใหญ่ จนถึงการวิเคราะห์ลำดับเบสในห้องปฏิบัติการขนาดเล็ก



**รูปที่ 8** (a) เทคโนโลยีการวิเคราะห์ลำดับเบสโดยนำสายดีเอ็นเอผ่านรูขนาดเส้นผ่านศูนย์กลางระดับนาโนเมตรหรือนาโนพอร์ (nanopore) ของ Oxford Nanopore Technologies<sup>®</sup> โมเลกุลดีเอ็นเอสามารถเคลื่อนผ่านรูนาโนพอร์ผ่านเมมเบรนโดยใช้ความเข้มข้นของไอออนที่ต่างกัน (b) เทคโนโลยีดีเอ็นเอทรานซิสเตอร์ (DNA transistor) ของ IBM สามารถอ่านเบสแต่ละตัวบนดีเอ็นเอสายเดี่ยวที่เคลื่อนผ่านนาโนพอร์ชนิดวัสดุที่เป็นของแข็งแล้วตรวจสัญญาณกระแสไฟฟ้าที่เกิดการเปลี่ยนแปลง [23]

#### 4. สรุป

ตลอดหลายปีที่ผ่านมา SGSTs มีการพัฒนาและก้าวหน้าอย่างชัดเจน การเกิดขึ้นใหม่ของรูปแบบการวิเคราะห์ลำดับเบสต่าง ๆ ไม่ว่าจะเป็น 454/Roche, Illumina/solexa, SOLiD/ABI หรือ Ion Torrent/Life technology ได้รับการยอมรับอย่างรวดเร็ว และถูกนำไปใช้อย่างกว้างขวางตามวัตถุประสงค์ของงานวิจัยสำหรับใช้วิเคราะห์ลำดับเบสจีโนม ช่วยในการค้นหา SNPs ด้วยต้นทุนค่าใช้จ่ายใน

การวิเคราะห์ลำดับเบสที่ลดลงอย่างรวดเร็ว ช่วยให้งานวิจัยในระดับเล็กสามารถเข้าถึงเทคโนโลยีดังกล่าวได้ จากที่ก่อนหน้านี้สามารถเข้าถึงได้เฉพาะงานวิจัยในระดับใหญ่ที่มีเงินทุนสูง อย่างไรก็ตามการเกิดขึ้นของ TGSTs จะยิ่งทำให้งานวิเคราะห์ลำดับเบสมีความก้าวหน้าอย่างมาก จะสามารถวิเคราะห์ลำดับเบสมนุษย์ทั้งจีโนมในเวลาหนึ่งวันด้วยค่าใช้จ่ายที่ถูกลงอย่างมาก นอกจากนี้เทคโนโลยีการวิเคราะห์ลำดับเบสยุคที่สามยังสามารถประยุกต์ใช้งานได้อย่างหลากหลาย

หลายมากขึ้นนอกจากการวิเคราะห์ลำดับเบสดีเอ็นเอ เช่นการวิเคราะห์รูปแบบการเติมหมู่เมทิลของดีเอ็นเอ การวิเคราะห์ลำดับเบสอาร์เอ็นเอเพื่อศึกษาทรานสคริปโตม การศึกษากระบวนการแปลรหัสโดยไรโบโซม การวิเคราะห์เหล่านี้จะช่วยให้การศึกษาระบบที่ซับซ้อนของสิ่งมีชีวิตมีความแม่นยำมากขึ้น ซึ่งเกิดประโยชน์อย่างยิ่งในการประยุกต์ใช้ในด้าน การแพทย์ ทั้งการวินิจฉัยโรค และการพัฒนาวิธีการรักษาโรคที่เกิดจากความผิดปกติทางพันธุกรรม รวมทั้งการประยุกต์ใช้ในการเกษตร ซึ่งคาดว่าจะสามารถช่วยให้เทคโนโลยีทางการแพทย์ และการเกษตรเจริญก้าวหน้าได้อย่างรวดเร็วแน่นอน

## 5. รายการอ้างอิง

- [1] Sanger, F., Nicklen, S. and Coulson, A.R., 1977, DNA sequencing with chain-terminating inhibitors, *Proc. Nat. Acad. Sci. USA*. 74: 5463-5467.
- [2] Smith, L.M., Sanders, J.Z., Kaiser, R.J., Hughes, P., Dodd, C., Connell, C.R., Heiner, C., Kent, S.B. and Hood, L.E., 1986, Fluorescence detection in automated DNA sequence analysis, *Nature* 321: 674-679.
- [3] Margulies, M., Egholm, M., Altman, W.E., *et al.*, 2005, Genome sequencing in microfabricated high-density picolitre reactors, *Nature* 437: 376-380.
- [4] Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlen, M. and Nyren, P., 1996, Real-time DNA sequencing using detection of pyrophosphate release, *Anal. Biochem*, 242: 84-89.
- [5] Son, M.S. and Taylor, R.K., 2011, Preparing DNA Libraries for Multiplexed Paired-End Deep Sequencing for Illumina GA Sequencers, *Current protocols in microbiology*, CHAPTER, Unit1E.4-Unit1E.4.
- [6] Mardis, E.R., 2013, Next-generation sequencing platforms, *Ann. Rev. Anal. Chem.* 6: 287-303.
- [7] Mardis, E.R., 2008, Next-generation DNA sequencing methods, *Ann. Rev. Genomics Hum. Genet.* 9: 387-402.
- [8] Brown, S.M., 2013, Next-generation DNA Sequencing Informatics, Cold Spring Harbor Laboratory Press, 241 p.
- [9] Ansorge, W.J., 2009, Next-generation DNA sequencing techniques, *N. Biotech.* 25: 195-203.
- [10] Merriman, B. and Rothberg, J.M., 2012, Progress in ion torrent semiconductor chip based sequencing, *Electrophoresis* 33: 3397-3417.
- [11] Rothberg, J.M., Hinz, W., Rearick, T.M., *et al.*, 2011, An integrated semiconductor device enabling non-optical genome sequencing, *Nature* 475: 348-352.
- [12] Eid, J., Fehr, A., Gray, J., *et al.*, 2009, Real-time DNA sequencing from single polymerase molecules. *Science* 323: 133-138.
- [13] Levene, M.J., Korch, J., Turner, S.W., *et al.*, 2003, Zero-mode waveguides for single-molecule analysis at high concentrations. *Science* 299: 682-686.

- [14] Flusberg, B.A., Webster, D.R., Lee, J.H., Travers, K.J., Olivares, E.C., Clark, T.A., Korlach, J. and Turner, S.W., 2010, Direct detection of DNA methylation during single-molecule, real-time sequencing, *Nat. Methods* 7: 461-465.
- [15] Uemura, S., Aitken, C.E., Korlach, J., Flusberg, B.A., Turner, S.W. and Puglisi, J.D., 2010, Real-time tRNA transit on single translating ribosomes at codon resolution, *Nature* 464: 1012-1017.
- [16] Kasianowicz, J.J., Brandin, E., Branton, D. and Deamer, D.W., 1996, Characterization of individual polynucleotide molecules using a membrane channel, *Proc. Nat. Acad. Sci. USA*. 93: 13770-13773.
- [17] Howorka, S., Cheley, S. and Bayley, H., 2001, Sequence-specific detection of individual DNA strands using engineered nanopores, *Nat. Biotech.* 19: 636-639.
- [18] Garaj, S., Hubbard, W., Reina, A., *et al.*, 2010, Graphene as a subnanometre trans-electrode membrane, *Nature* 467: 190-193.
- [19] Kawano, R., Schibel, A.E., Cauley, C., White, H.S., 2009, Controlling the translocation of single-stranded DNA through alpha-hemolysin ion channels using viscosity, *Langmuir* 25: 1233-1237.
- [20] Lieberman, K.R., Cherf, G.M., Doody, M.J., *et al.*, 2010, Processive replication of single DNA molecules in a nanopore catalyzed by phi29 DNA polymerase, *J. Amer. Chem. Soc.* 132: 17961-17972.
- [21] Clarke, J., Wu, H.C., Jayasinghe, L., *et al.*, 2009, Continuous base identification for single-molecule nanopore DNA sequencing, *Nat. Nanotech.* 4: 265-270.
- [22] McNally, B., Singer, A., Yu, Z., *et al.*, 2000, Optical recognition of converted DNA nucleotides for single molecule DNA sequencing using nanopore arrays, *Nano Lett.* 10: 2237-2244.
- [23] Schadt, E.E., Turnerm, S. and Kasarskis, A., 2010, A window into third-generation sequencing *Hum. Mol. Gen.* 19: 227-240.