

วิวัฒนาการของสายพันธุ์ไวรัสก่อโรคโควิด-19 และกระบวนการทางชีวสารสนเทศในการตรวจสอบสายพันธุ์

Evolution and Bioinformatic Analysis of the Whole Genome Sequences of SARS-CoV-2

ธนัญฐ์ ศิลา, ทน., วท.บ.¹, วิศลย์ เหล่าเจริญสุข, พ.บ.^{2,3},
คมวิทย์ สุรชาติ, ประ.ด.², สมนรพรพร สุระสมบัติพัฒนา, ทน., วท.บ.¹,
วันท แซ่ลิม, วท.บ.³, สุรศักดิ์ สังขทัต ณ อยุธา, พ.บ., ประ.ด.^{3,4}
Thanit Sila, MT, B.Sc.¹, Wison Laochareonsuk, M.D.^{2,3},
Komwit Surachart, Ph.D.², Samonrapat Surasombatpattana, MT, Ph.D.¹,
Wanut Saelim, Bsc.³, Surasak Sangkhathat, M.D., Ph.D.^{3,4}

¹สาขาวิชาพยาธิวิทยา คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์ อำเภอหาดใหญ่ จังหวัดสงขลา 90110

²สาขาวิชาชีวเวชศาสตร์และวิศวกรรมชีวการแพทย์ คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์ อำเภอหาดใหญ่ จังหวัดสงขลา 90110

³ศูนย์วิจัยเวชศาสตร์ปริวรรต คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์ อำเภอหาดใหญ่ จังหวัดสงขลา 90110

⁴สาขาวิชาศัลยศาสตร์ คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์ อำเภอหาดใหญ่ จังหวัดสงขลา 90110

¹Department of Pathology, Faculty of Medicine, Prince of Songkla University, Hat Yai, Songkhla 90110, Thailand.

²Department of Biomedical Sciences & Biomedical Engineering, Faculty of Medicine, Prince of Songkla University, Hat Yai, Songkhla 90110, Thailand.

³Translational Medicine Research Center, Faculty of Medicine, Prince of Songkla University, Hat Yai, Songkhla 90110, Thailand.

⁴Department of Surgery, Faculty of Medicine, Prince of Songkla University, Hat Yai, Songkhla 90110, Thailand.

Corresponding author: Thanit Sila, MT, B.Sc.

Department of Pathology, Faculty of Medicine, Prince of Songkla University,
Hat Yai, Songkhla 90110, Thailand.

E-mail: thanit_miaw@hotmail.com

doi: 10.31584/psumj.2022254705

<https://he01.tci-thaijo.org/index.php/PSUMJ/>

PSU Med J 2022;2(2):85-98

Received 8 February 2022

Revised 6 March 2022

Accepted 6 March 2022

Published online 19 May 2022

บทคัดย่อ

โรคติดเชื้อไวรัสโคโรนา 2019 (COVID-19) ซึ่งเกิดจากเชื้อไวรัส (SARS-CoV-2) มีการแพร่ระบาดของเชื้ออย่างกว้างขวางไปทั่วโลก อัตราการแพร่กระจายของเชื้อที่มากขึ้นก็จะส่งผลให้เกิดการกลายพันธุ์ที่เร็วขึ้น มีการพบการกลายพันธุ์ที่ตำแหน่งโปรตีนหนาม (spike protein) ของ SARS-CoV-2 ที่ส่งผลให้เกิดการเปลี่ยนแปลงในการเข้าจับกับ ACE2 receptor บนเซลล์บุผิวทางเดินหายใจของมนุษย์ซึ่งเป็นปัจจัยสำคัญในการติดเชื้อ ความรุนแรง และความเร็วในการแพร่ระบาดอย่างมาก ปัจจุบันนักวิทยาศาสตร์ได้ตระหนักและให้ความสำคัญกับสายพันธุ์ที่น่ากังวลที่กำลังระบาดและส่งผลกระทบต่อวงกว้างได้แก่ สายพันธุ์แอลฟา สายพันธุ์เบต้า สายพันธุ์แกมมา สายพันธุ์เดลต้าและสายพันธุ์โอไมครอน ดังนั้นการเฝ้าระวังและลดการแพร่ระบาดของโรคติดเชื้อไวรัสโคโรนา 2019 จึงจำเป็นต้องเข้าใจโครงสร้าง การทำงานและการกลายพันธุ์ของไวรัส กระบวนการทางชีวสารสนเทศเป็นเครื่องมือหนึ่งที่ใช้ตรวจสอบสายพันธุ์ในรูปแบบข้อมูลรหัสพันธุกรรมซึ่งได้มาโดยกระบวนการทางอณูชีวโมเลกุล ไม่ว่าจะเป็นวิธี RT-PCR, NGS, Mass Array และอื่น ๆ แม้ว่าการควบคุมการระบาดของโรคติดเชื้อไวรัสโคโรนา 2019 ยังไม่มีที่สิ้นสุดในปัจจุบัน การรู้เท่าทันการเปลี่ยนแปลงของเชื้อไวรัสยังคงวิวัฒนาการอย่างต่อเนื่องและเป็นไปในทิศทางซึ่งทำให้ไวรัสมีศักยภาพสูงขึ้นในการติดเชื้อและหลบเลี่ยงภูมิคุ้มกันเป็นสิ่งสำคัญ

คำสำคัญ: จีโนมของไวรัส; โรคโควิด 2019; ไวรัส SAR-CoV-2

ABSTRACT

Coronavirus Disease 2019 (COVID-19) disease, caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) virus, has been spreading as a pandemic and health catastrophe. The increase in the rate of infection influenced the mutation rate to increase greatly. Mutations found in the spike protein site of SARS-CoV-2 may result in changes in the binding of the angiotensin-converting enzyme 2 (ACE2) receptor in the human respiratory mucosa, which is a significant step in infection, severity, and speed of the epidemic. Scientists are now recognizing and focusing on variant of concern (VOC) that are currently spreading and affecting: Alpha, Beta, Gamma, Delta, and Omicron. Therefore, to monitor and mitigate the spread of coronavirus-19, it is imperative to understand the structure, functions, and mutation of the virus. Bioinformatics is one tool used to determine strains in the genetic code data obtained by molecular biomolecular processes such as real-time polymerase chain reaction (RT-PCR), next generation sequencing (NGS), Mass Array and others. As the COVID-19 pandemic seems not to end soon, viruses continue to evolve, and this has given them a greater potential for infection and evasion of immunity, monitoring of viral genome is crucial.

Keywords: Covid-19; SARS-CoV-2; Virus Genome

บทนำ

ในขณะที่โลกกำลังเผชิญกำลังเพื่อเอาชนะการระบาดของโรคติดเชื้อไวรัสโคโรนา 2019 ด้วยมาตรการเว้นระยะห่างทางสังคม การจำกัดการเดินทาง วัคซีน ตลอดจนการค้นคว้าหายาด้านไวรัสที่มีประสิทธิภาพ เชื้อไวรัสก่อโรคติดเชื้อไวรัสโคโรนา 2019 เองก็มีศักยภาพวิวัฒนาการสายพันธุ์จนกระทั่งมีความสามารถในการเข้าติดเชื้อและแพร่กระจายได้รวดเร็วขึ้น¹ กลยุทธ์ในการอยู่เหนือไวรัสและควบคุมการระบาด

จะเป็นไปได้ยากหากไม่สามารถตรวจสอบทราบสายพันธุ์ของไวรัส เพื่อใช้ข้อมูลในการสืบสวนการระบาดและเฝ้าระวังป้องกันการระบาดของสายพันธุ์ซึ่งมีหลักฐานของความน่ากังวล ในปัจจุบันหลายองค์กรทางสาธารณสุขในโลกรวมทั้งประเทศไทยกำลังระดมสรรพกำลังในการศึกษาพันธุกรรมของเชื้อไวรัสและเผยแพร่ในฐานข้อมูลสาธารณะเพื่อใช้ประโยชน์ร่วมกัน ในลักษณะการจัดการข้อมูลความรู้เพื่อต่อสู้การระบาด

(infodemics)²⁻⁴ ข้อมูลการระบาดของสายพันธุ์ไวรัสซึ่งมีการกลายพันธุ์ที่นำกังวล ถูกเฝ้าระวังและป้องกันไปยังการยกระดับความเข้มข้นของมาตรการป้องกัน⁵

บทความนี้มีเป้าหมายเพื่อที่จะสร้างความเข้าใจในข้อมูลพื้นฐานของเชื้อไวรัส SARS-CoV-2 การกลายพันธุ์ ตลอดจนผลกระทบของการกลายพันธุ์ที่มีต่อความรุนแรงของโรคและความรุนแรงของการระบาด และจะได้นำมาสู่เทคนิควิธีการทางห้องปฏิบัติการและกระบวนการทางชีวสารสนเทศซึ่งใช้ในการตรวจสอบลำดับสารพันธุกรรมของไวรัสในปัจจุบันโดยสังเขป

ไวรัสก่อโรคโควิด-19

ในเดือนธันวาคม พ.ศ. 2562 มีรายงานการติดเชื้อของปอดที่ไม่ทราบเหตุ (pneumonia of an unknown etiology) เกิดขึ้นที่เมืองอู่ฮั่น (Wuhan) มณฑลหูเป่ย์ (Hubei) ประเทศจีน การสืบสวนการระบาดพบว่าผู้ป่วย 4 รายแรกมีประวัติเดินทางไปยังตลาดค้าส่งอาหารทะเลหัวหนาน (South China Seafood Wholesale Market)⁶ ไวรัสซึ่งแยกได้จากเซลล์บุผิวหลอดลมของผู้ป่วยอยู่ในตระกูลเดียวกับไวรัส Corona ในเบื้องต้นจึงได้รับการตั้งชื่อชั่วคราวโดยองค์การอนามัยโลกว่าเป็น 2019-nCoV⁷ ข้อมูลจากการศึกษาลำดับสารพันธุกรรมของไวรัสระยะถัดมาซึ่งพบว่าไวรัสชนิดนี้อยู่ในสายพันธุ์ (lineage) เดียวกับไวรัสก่อโรค severe acute respiratory syndrome กล่าวคือ SARS-CoV ซึ่งระบาดหนักในประเทศจีนและเอเชียตะวันออกเฉียงใต้ 10 ปีที่ผ่านมา ในระยะถัดมา International Committee for the Classification of Viruses จึงพิจารณาให้ชื่อไวรัสสายพันธุ์ใหม่นี้ว่า SARS-CoV-2 ในวันที่ 11 กุมภาพันธ์ พ.ศ.2563 และในขณะเดียวกัน องค์การอนามัยโลกให้ชื่อโรคซึ่งเกิดจากการติดเชื้อ SARS-CoV-2 ว่า coronavirus disease 2019 หรือโรคติดเชื้อไวรัสโควิด 2019 (COVID-19)⁸

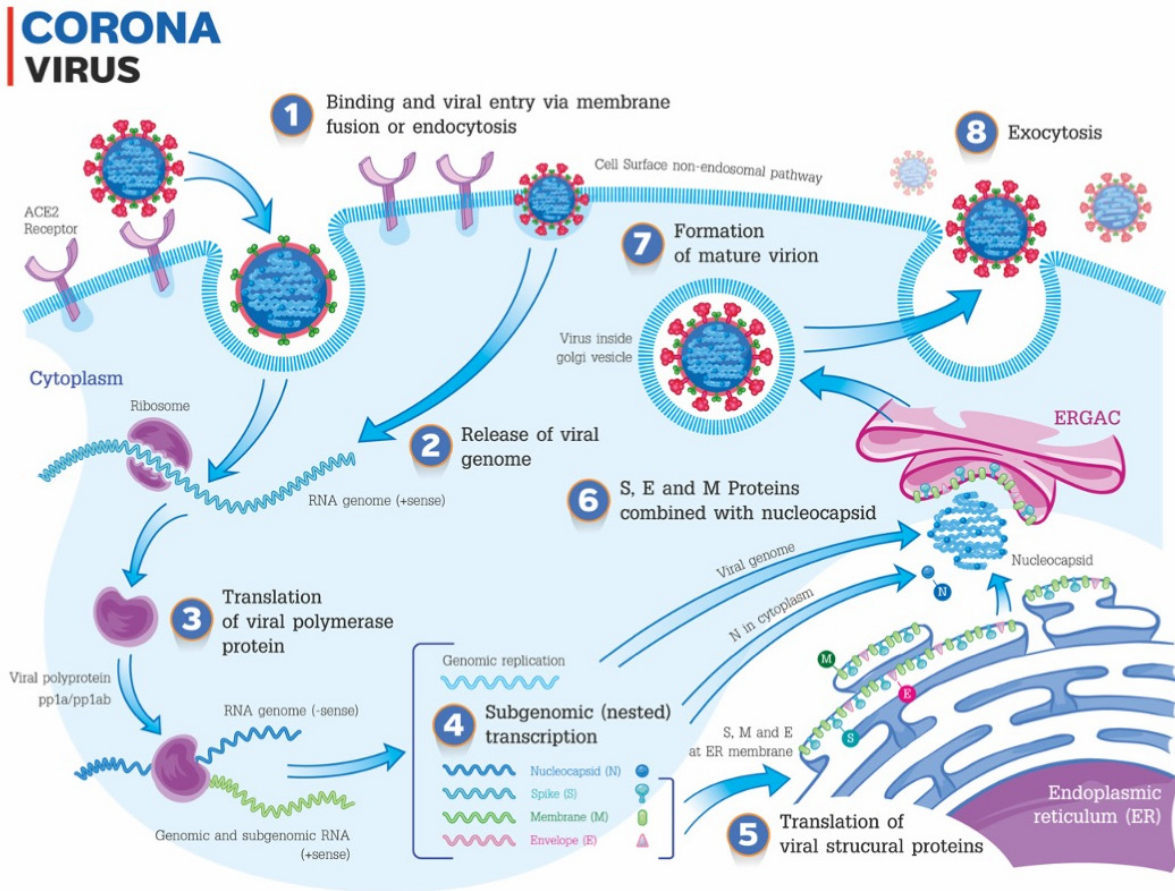
ช่องทางการระบาดหลักของโควิด 19 คือการสัมผัสทางร่างกายและทางเดินหายใจ แม้จะมีหลักฐานว่าไวรัสอาจปนเปื้อนอยู่ในอุจจาระแต่ก็ยังไม่เคยมีรายผู้ป่วยที่รับเชื้อจากช่องทางดังกล่าว การระบาดของโรคกระจายไปอย่างกว้างขวางนับถึงวันที่เตรียมต้นฉบับนี้ในปลายเดือนตุลาคม พ.ศ.2564 มีรายงานผู้ติดเชื้อทั่วโลกกว่า 250 ล้านราย และมีผู้เสียชีวิตจากการติดเชื้อทั่วโลกกว่า 5 ล้านราย ในขณะที่ในประเทศไทยมีผู้ติดเชื้อที่ได้รับการรายงานราว 1.8 ล้านรายและมีการเสียชีวิต 1.8 หมื่นราย⁹ ปัจจัยสำคัญประการหนึ่งซึ่งทำให้การ

ควบคุมการระบาดเป็นไปอย่างยากลำบากคือการกลายพันธุ์ของเชื้อ SARS-CoV-2 ซึ่งส่งผลให้เชื้อมีศักยภาพในการเกาะติดเซลล์บุผิวและสามารถหลบหลีกระบบภูมิคุ้มกันที่ได้จากการฉีดวัคซีน⁹

ข้อมูลพื้นฐานของไวรัส SARS-CoV-2

ไวรัส SARS-CoV-2 อยู่ในแฟมิลี Coronaviridae จีนัส Betacoronavirus และร่วมอยู่ในจีนัสย่อย Sarbecovirus (lineage B) ซึ่ง SARS-CoV-2 เป็นตัวล่าสุดหรือตัวที่ 7 ของไวรัสในกลุ่ม Betacoronavirus¹⁰ ซึ่งก่อโรคในคน โดย 6 ตัวที่ค้นพบก่อนหน้านี้ได้แก่ ไวรัส HCoV-229E, HCoV-NL63, HCoV-OC43 และ HCoV-HKU1 ซึ่งสัมพันธ์กับไข้หวัดตามฤดูกาล ไวรัส SARS-CoV ซึ่งก่อให้เกิดการระบาดเป็นวงกว้างในปี พ.ศ.2545-2547 และ MERS-CoV ซึ่งระบาดในประเทศแถบตะวันออกกลางในปี พ.ศ.2555¹¹ สมบัติพื้นฐานของไวรัสในกลุ่มนี้คือเป็น single-stranded RNA สายบวก ซึ่งมีขนาดความยาวของจีโนมตั้งแต่ 26-32 กิโลเบส ลักษณะทางสัณฐานวิทยาของโปรตีนโครงสร้างไวรัสในตระกูล Corona มีลักษณะทรงกลม มีหนามคล้ายรัศมีซึ่งแผ่ออกจากดวงอาทิตย์ (รูปที่ 1) เข้าบ้านซึ่งเป็นแหล่งเก็บเชื้อตามธรรมชาติของไวรัสกลุ่มนี้คือค้างคาว อย่างไรก็ตามไวรัสสามารถแพร่ผ่านสัตว์ที่เป็นเจ้าบ้านตัวกลาง (intermediate host) ได้แก่ อูฐ ตัวนิ่ม (pangolin) รวมทั้งคนเอง กระทั่งก่อการระบาดและพยาธิสภาพของทางเดินหายใจในคน

ไวรัส SARS-CoV-2 มีความยาวสารพันธุกรรมตลอดจีโนมทั้งสิ้น 29,903 เบส ซึ่งมีลำดับสารพันธุกรรมตรงกับ SARS-CoV ร้อยละ 79.5⁷ และตรงกับสายพันธุ์ RatG13 ซึ่งระบาดในค้างคาวร้อยละ 96¹¹ จีโนมของ SARS-CoV-2 ประกอบด้วย open reading frame (ORF) อย่างน้อย 14 หน่วย ที่เข้ารหัสสำหรับโปรตีน 27 ตัว⁶ แบ่งออกเป็นโปรตีนโครงสร้าง (structural protein) และส่วนที่เข้ารหัสสำหรับโปรตีนที่นอกเหนือจากโปรตีนโครงสร้าง (non-structural protein)¹² โดยโปรตีนโครงสร้างของ SARS-CoV-2 ประกอบด้วย 4 ส่วน (รูปที่ 1) คือ โปรตีนหนาม (spike protein) โปรตีนเปลือก (envelope protein) โปรตีนผิว (membrane protein) โดยทั้งสามส่วนนี้ประกอบขึ้นเป็นผิวนอกของทรงกลมซึ่งทำหน้าที่ห่อหุ้มสารพันธุกรรมของไวรัสและโปรตีนนิวคลีโอแคปซิด (nucleocapsid protein) ซึ่งอยู่ภายในและทำหน้าที่ประคองเพื่อรักษาเสถียรภาพของสาร



รูปที่ 1 วงจรชีวิตของไวรัส SARS-CoV-2 ซึ่งก่อโรคโควิด 19

- 1.) การเข้าสู่เซลล์เจ้าบ้านโดยการเกาะจับกับตัวรับ ACE2
- 2.) ปลดปล่อยสารพันธุกรรมออกจากโปรตีนเปลือก
- 3.) ถอดรหัสและสร้างเอนไซม์ที่จำเป็นต่อการจำลองแบบ
- 4.-5.) จำลองแบบส่วนประกอบของไวรัส
- 6.-7.) ประกอบร่างใหม่จากองค์ประกอบที่สร้างขึ้น
- 8.) ออกจากเซลล์เจ้าบ้านโดยกระบวนการ exocytosis

พันธุกรรมไวรัส¹³ ส่วนของจีโนมนอกเหนือจากที่เข้ารหัสโปรตีนโครงสร้างกินความยาวราว 2 ใน 3 ของความยาวทั้งหมดและอยู่ทางปลายฝั่ง 5' ส่วนดังกล่าวเป็นส่วน ORF ซึ่งเข้ารหัสสำหรับเอนไซม์ซึ่งมีความจำเป็นต่อการสร้างสำเนาสารพันธุกรรม อาทิเช่น เอนไซม์ proteases, RNA polymerase, helicase และ ribonuclease¹⁴

ความสำคัญของโปรตีนโครงสร้างที่มีต่อพยาธิสรีรวิทยาของการติดเชื้อไวรัส SARS-CoV-2

ไวรัส SARS-CoV-1 และ SARS-CoV-2 ใช้กลไกร่วมในการยึดเกาะผิวเซลล์และแทรกตัวเข้าเซลล์เจ้าบ้านโดยการจับกันระหว่างโปรตีนหนามบนเปลือกไวรัสกับโปรตีน

ตัวรับ angiotensin-converting enzyme 2 (ACE2) บนเยื่อหุ้มเซลล์เจ้าบ้าน¹⁵ โปรตีนตัวรับ ACE2 เป็นโปรตีนคล้าย (analogue) ของ angiotensin-converting enzyme โดยในมนุษย์พบการแสดงออกของ ACE2 มากที่สุดในทางเดินหายใจ นอกจากนี้ยังพบได้ที่ปอด หัวใจ และไต โดยโปรตีนหนามเป็นกลัยโคโปรตีนสามกลีบ (trimeric) ขนาด 180-200 kDa ประกอบด้วยกรดอะมิโน 1,273 ตัว แบ่งส่วนย่อยได้ 3 โดเมน (รูปที่ 1) คือ N-terminus domain (NTD), S1 และ S2 เมื่อเข้าสู่ร่างกายมนุษย์ โปรตีนหนามถูกกระตุ้นและตัดโดยเอนไซม์ furin และ transmembrane serine protease TMPRSS2 ออกเป็น 2 หน่วยย่อย S1 และ S2¹⁶ จากนั้น receptor binding domain (RBD) บน

หน่วยย่อย S1 ของโปรตีนหนามทำปฏิกิริยากับ ACE2 เพื่อยึดเกาะบนเยื่อหุ้มเซลล์ ในขณะที่บนหน่วยย่อย S2 โดเมน fusion peptide (FP), heptapeptide repeat sequence (HR1), HR2, โดเมน transmembrane (TM) และโดเมน CT domain (CT) ทำหน้าที่หลอมเปลือกหุ้มไวรัสเข้ากับเยื่อหุ้มเซลล์เพื่อนำเข้าสารพันธุกรรมไวรัส โปรตีนหนามเป็นส่วนของไวรัสซึ่งถูกนำไปใช้ในการออกแบบยามากที่สุด¹⁴

เมื่อไวรัสประสบความสำเร็จในการเข้าสู่เซลล์เจ้าบ้าน เปลือกหุ้มจะถูกสลายและสารพันธุกรรมถูกปลดปล่อยออกสู่ไซโตพลาสซึม ถอดรหัสเป็น mRNA และโปรตีนที่เกี่ยวข้องพร้อมกับการขยายสำเนาสารพันธุกรรม เพื่อประกอบร่างเป็นไวรัสหน่วยใหม่และถูกปล่อยออกนอกเซลล์

การกลายพันธุ์ของ SARS-CoV-2 และการจำแนกสายพันธุ์

เนื่องจากไวรัสเป็นสิ่งมีชีวิตที่ขยายพันธุ์ได้ในเวลาอันสั้น และแพร่พันธุ์อย่างรวดเร็ว การเปลี่ยนแปลงของสารพันธุกรรมบนจีโนมไวรัสอันเนื่องมาจากความผิดพลาดในขณะจำลองแบบ (replication error) ซึ่งเกิดขึ้นแบบสุ่มเป็นปกติ อัตราการกลายพันธุ์ของ SARS-CoV-2 ได้รับการประมาณว่าอยู่ที่ 10^{-6} ครั้งต่อเบสต่อรอบการจำลองแบบ ซึ่งน้อยกว่าอัตราการกลายพันธุ์ที่พบในไวรัสไข้หวัดใหญ่อันเนื่องมาจาก SARS-CoV-2 มีกลไกแก้ไขความผิดพลาดใน-ขณะจำลองแบบโดยใช้เอนไซม์ Exonuclease (ExoN)¹¹ ความเร็วในการกลายพันธุ์จะสูงขึ้นในพื้นที่ซึ่งมีความชุกของการระบาดสูง การจำแนกสายพันธุ์ของไวรัส SARS-CoV-2 มีหลายระบบ ไวรัส 2 สายพันธุ์แรกที่แยกได้จากเมืองอู่ฮั่นคือ lineage B (Wuhan-Hu-1) ซึ่งได้จากตัวอย่างเก็บเมื่อวันที่ 26 ธันวาคม พ.ศ. 2562 และ lineage A (Wuhan/WH04/2020 หรือ EPI_ISL_406801) ซึ่งได้จากตัวอย่างเก็บเมื่อวันที่ 5 มกราคม พ.ศ. 2563¹⁷ สายพันธุ์ใหม่ที่เกิดขึ้นอาจมีความแตกต่างของสารพันธุกรรมจากไวรัสที่ระบาดก่อนหน้านี้เพียงตำแหน่งเดียวหรือหลายตำแหน่ง และเมื่อเกิดขึ้น การระบาดของสายพันธุ์นั้นอาจชะลอกระทั่งสิ้นสุดลงหรือขยายวงออกไปขึ้นกับศักยภาพในการแพร่พันธุ์ของสายพันธุ์กลายพันธุ์นั้น ซึ่งถูกกำหนดโดยตำแหน่งและชนิดของกรดอะมิโนที่มีการเปลี่ยนแปลง โดยเฉพาะอย่างยิ่งบนโดเมนที่เข้ารหัสโปรตีนหนาม

การให้ชื่อสายพันธุ์ได้จากการศึกษาข้อมูลลำดับเบสของไวรัสซึ่งได้รับการศึกษาจากหลากหลายพื้นที่ซึ่งได้รับการเผยแพร่บนเว็บไซต์ของฐานข้อมูล Global Initiative on Sharing All Influenza Data (GISAID)¹⁸ การเปรียบเทียบและให้ชื่อ (nomenclature) ของสายพันธุ์ไวรัสมีหลายระบบ การให้ชื่อสายพันธุ์ในระยะเริ่มต้นใช้ระบบ GISAID clade ซึ่งจัดกลุ่มไวรัสเป็นสาขาหลักใหญ่ (major clades) 8 กลุ่มโดยใช้เครื่องหมายการกลายพันธุ์ซึ่งระบุการแตกสาขาออกจากสายพันธุ์ S และ L ซึ่งเป็นสายพันธุ์เริ่มต้น ออกสู่สายพันธุ์ V, G, GK, GH, GR และ GV วิเคราะห์จากข้อมูลการตรวจสอบลำดับเบสตลอดจีโนมและเปรียบเทียบโดยใช้การคำนวณแบบ phylogenetic tree อย่างไรก็ตามการจำแนกสาขาหลักตามแบบของ GISAID clade ไม่ตอบสนองต่อการให้รายละเอียดของการเปลี่ยนแปลงที่เกิดขึ้นที่ละน้อยแต่มีพลวัตการเปลี่ยนแปลงที่รวดเร็ว จึงมีการกำหนดรูปแบบการให้ชื่อสายพันธุ์แบบใช้ตัวเลขต่อท้ายสายพันธุ์หลัก เช่น B.1 เมื่อมีหลักฐานว่าการระบาดหนักในทีใดทีหนึ่งก่อให้เกิดการแยกสายพันธุ์ (ตารางที่ 1) และมีผู้ติดตามของสายพันธุ์ที่เกิดขึ้นใหม่นั้นมากพอที่จะมีโอกาสระบาดต่อเนื่องไปยังตำแหน่งทางภูมิศาสตร์อื่น และเข้าเกณฑ์ดังต่อไปนี้ ก. มีการเปลี่ยนแปลงของลำดับเบสอย่างน้อย 1 ตำแหน่งจากสายพันธุ์ต้นสาขาหลัก (ancestral lineage) ข. พบสายพันธุ์ใหม่ดังกล่าวอย่างน้อย 5 จีโนมขึ้นไป ค. สายพันธุ์ใหม่จะต้องมีตำแหน่งที่มีการกลายพันธุ์ร่วมอย่างน้อย 1 ตำแหน่ง และ ง. เมื่อทำการสุ่มซ้ำ (bootstrap) ยังคงพบจีโนมที่เข้ากับสายพันธุ์ใหม่ (lineage-defining node) ได้มากกว่าร้อยละ 70 โดยสายพันธุ์ใหม่ที่เกิดขึ้นเมื่อมีการระบาดต่อเนื่องก็มีโอกาสที่จะเป็นสายพันธุ์ต้นสาขาหลักให้แตกเป็นสายพันธุ์ใหม่ต่อไปอีก¹⁷ ซึ่งจะทำให้มีการเพิ่มจุดทศนิยมและตัวเลขใหม่จนกระทั่งครบ 3 จุดจึงจะมีการขึ้นอักษรใหม่ เช่น B.1.617 (สายพันธุ์อินเดีย) แตกสาขาหลักมาจากสายพันธุ์ B.1 และสายพันธุ์ที่ระบาดอย่างรุนแรงในช่วงไตรมาสที่สองของปี พ.ศ. 2564 คือ B.1.617.2 (เดลต้า) วิวัฒนาการมาจาก B.1.617 อีกทอดหนึ่ง การเกิดสายพันธุ์มักสัมพันธ์กับการระบาดใหญ่ในแต่ละพื้นที่ทางภูมิศาสตร์ จึงมีการเรียกสายพันธุ์ตามพื้นที่การระบาด เช่น B.1.1.7 เป็นสายพันธุ์ซึ่งเคยเรียกว่าสายพันธุ์อังกฤษ ระบบการให้ชื่อสายพันธุ์ลักษณะนี้คือระบบ Phylogenetic Assignment of Named Global Outbreak LINEages (Pango lineages หรือ PANGOLIN)

สายพันธุ์ที่น่าสนใจ (Variants of Interest, VOI) และสายพันธุ์ที่น่ากังวล (Variants of Concern, VOC)

การกลายพันธุ์ของไวรัสเกิดขึ้นแบบสุ่มและถูกเร่งโดยความรุนแรงของการระบาด การกลายพันธุ์ส่วนใหญ่ไม่ส่งผลกระทบต่อความรุนแรงของโรคหรือศักยภาพในการแพร่พันธุ์ต่อเนื้อเยื่อ อย่างไรก็ตาม การกลายพันธุ์ที่เกิดกับโดเมนหรือตำแหน่งที่มีอิทธิพลต่อการเข้าสู่ร่างกายหรือการหลบหลีกระบบภูมิคุ้มกัน อาจส่งผลให้การระบาดทวีความรุนแรงและอัตราการป่วยตายสูงขึ้นดังเช่นที่เกิดกับสายพันธุ์ B.1.167.2 ที่กำเนิดขึ้นจากประเทศอินเดีย หรือที่รู้จักกันในนามสายพันธุ์เดลต้า ตามการให้นามไวรัสขององค์การอนามัยโลก ขณะเตรียมต้นฉบับบทความนี้ในเดือนตุลาคม พ.ศ. 2564 สายพันธุ์เดลต้าเป็นสายพันธุ์หลักของ SAR-CoV-2 ที่กำลังระบาดในประเทศไทย และกำลังมีรายงานการกลายพันธุ์ต่อเนื่องไปยัง AY.1 และ AY.2

คณะกรรมการติดตามวิวัฒนาการการระบาดขององค์การอนามัยโลกติดตามและกำหนดชื่อให้แก่ไวรัส 2 กลุ่มคือ VOI และ VOC โดยมีเกณฑ์วินิจฉัย คือ

สายพันธุ์ที่น่าสนใจ (VOI)

การกลายพันธุ์บนตำแหน่งซึ่งทราบโดยการทำนายหรือมีข้อมูลสนับสนุนว่า อาจก่อให้เกิดการเปลี่ยนแปลงในสมบัติของเชื้อ ได้แก่ การระบาด ความรุนแรงของโรค การหลบหลีกภูมิคุ้มกัน การหลบหลีกกระบวนการวินิจฉัยหรือการรักษา

เริ่มเกิดการระบาดของสายพันธุ์ดังกล่าวในหลายคลัสเตอร์ ในหลายประเทศ มีการเพิ่มขึ้นของความชุกและจำนวนผู้ป่วยตามระยะเวลาการระบาด หรือมีผลกระทบต่อ

เชิงระบาดวิทยาต้านอื่นซึ่งทำให้เชื่อได้ว่าจะมีผลกระทบต่อสาธารณสุขระดับโลก

เมื่อไม่นานมานี้มีสายพันธุ์ที่น่าสนใจ 2 ตัวได้แก่ Lambda (C.37 หรือ GR/452Q.V1) ซึ่งมีจุดเริ่มการระบาดจากประเทศเปรูเมื่อเดือนธันวาคม พ.ศ. 2563 และ Mu (B.1.621 หรือ GH) ซึ่งมีจุดเริ่มการระบาดจากประเทศโคลอมเบียในเดือน มกราคม พ.ศ. 2564 ซึ่งเกิดขึ้นในช่วงระหว่างที่กำลังเตรียมต้นฉบับในเดือนตุลาคม พ.ศ. 2564

สายพันธุ์ที่น่ากังวล (VOC)

หมายถึงสายพันธุ์ VOI เดิมซึ่งมีหลักฐานจากการศึกษาเปรียบเทียบ (comparative assessment) ชัดเจนว่า การกลายพันธุ์ทำให้เชื้อระบาดอย่างรวดเร็วขึ้น หรือเพิ่มความรุนแรงของโรคหรืออาการทางคลินิก ลดประสิทธิภาพของมาตรการควบคุมทางสังคมและระบบสาธารณสุข ลดการตอบสนองต่อยา วัคซีนหรือวิธีการวินิจฉัยโรคที่มีในขณะนั้น (ตารางที่ 1)

เมื่อ VOI/VOC เกิดขึ้น และถูกรายงานไปยังองค์การอนามัยโลก ประเทศซึ่งมีการระบาดของสายพันธุ์ที่น่าสนใจหรือน่ากังวลดังกล่าวจะได้รับการร้องขอให้ส่งข้อมูลลำดับเบสของไวรัสที่สมบูรณ์ขึ้นสู่ฐานข้อมูลสาธารณะ ได้แก่ GISAID และศึกษาการระบาดของสายพันธุ์ที่กำลังถูกจับตาดู และนอกจาก VOI/VOC องค์การอนามัยโลกยังติดตาม ไวรัสกลายพันธุ์จำนวนหนึ่งโดยยังไม่ให้ชื่อเฉพาะ เรียกว่ากลุ่ม Variants Under Monitoring หรือ VUM และไวรัสในกลุ่มนี้ส่วนหนึ่งถูกลดระดับมาจาก VOI จึงได้รับชื่อไปแล้วเช่น สายพันธุ์ Kappa หรือ B.1.617.1 ซึ่งเคยระบาดในประเทศอินเดีย

ตารางที่ 1 สายพันธุ์ที่น่ากังวล (variants of concern) ตามการประกาศขององค์การอนามัยโลก

ประเทศเริ่มระบาด	ชื่อสายพันธุ์ตามระบบการจัดจำแนก		
	WHO (วันที่ประกาศเป็น VOC)	PANGOLIN	GISAID claid
สหราชอาณาจักร (ก.ย. 2563)	Alpha (18 ธ.ค. 63)	B.1.1.7	GRY
แอฟริกาใต้ (พ.ค. 2563)	Beta (18 ธ.ค. 63)	B.1.351	GH/501Y.V2
บราซิล (พ.ย. 2563)	Gamma (11 ม.ค. 64)	P.1	GR/501Y.V3
อินเดีย (เม.ย. 2564)	Delta (11 พ.ค. 64)	B.1.617.2	G/478K.V1
แอฟริกาใต้ (พ.ย. 2564)	Omicron (26 พ.ย. 64)	B.1.1.529	GRA

WHO = World Health Organization; GISAID claid = Global Initiative on Sharing All Influenza Data

การจัดกลุ่ม SARS-CoV-2 เป็น VOC/VOI ตามระบบขององค์การอนามัยโลกใช้ชื่อกลุ่มและคำจำกัดความคล้ายกับการจัดกลุ่มไวรัสของศูนย์ควบคุมและป้องกันโรคแห่งประเทศสหรัฐอเมริกา (Center for Disease Control and Prevention, CDC) ซึ่งนอกจากจะแบ่ง SARS-CoV-2 ออกเป็น VOC, VOI แล้ว ยังมีกลุ่ม Variants Being Monitors (VBM) ซึ่งเป็นกลุ่มที่กำลังถูกติดตามเฝ้าระวังในประเทศสหรัฐอเมริกาปัจจุบัน โดย VOC ที่เหลืออยู่ในมุมมองของ CDC ในปัจจุบัน (มกราคม 2565) คือ B.1.617.2 หรือ สายพันธุ์ AY และ B.1.1.529 หรือสายพันธุ์ BA ส่วน VOC/VOI ตามการจำแนกขององค์การอนามัยโลกส่วนใหญ่อยู่ในกลุ่ม VBM และมีการนิยามไวรัสที่ก่อให้เกิดการระบาดที่รุนแรง (Variant of High Consequence-VOHC) ในสังคมที่ได้รับวัคซีนอย่างเต็มที่แล้ว ซึ่งในปัจจุบันยังไม่มีไวรัสที่เข้าเกณฑ์ดังกล่าว

การกลายพันธุ์ที่น่ากังวลบนสายพันธุ์ที่น่ากังวล สายพันธุ์อัลฟา (B.1.1.7)

สายพันธุ์อัลฟาเป็น VOC ที่เริ่มระบาดจากประเทศอังกฤษในเดือนกันยายน พ.ศ. 2563 และลุกลามไปยังทวีปยุโรป แคนาดาและสหรัฐอเมริกาในช่วงฤดูหนาว ความน่ากังวลของสายพันธุ์อัลฟาคือไวรัสตอบสนองต่อวัคซีน AstraZeneca ซึ่งเป็นวัคซีนหลักในแถบทวีปยุโรปเพียงร้อยละ 70 อย่างไรก็ตาม สายพันธุ์นี้ยังตอบสนองต่อวัคซีนกลุ่ม mRNA มากกว่าร้อยละ 90 นอกจากนี้สายพันธุ์อัลฟาเป็น VOC ตัวแรกที่ได้รับการพิสูจน์ว่ามีระบาดที่เร็วขึ้น

การกลายพันธุ์บนสายพันธุ์อัลฟาเกิดขึ้น 17 ตำแหน่ง (รูปที่ 2) ในจำนวนนี้ 8 ตำแหน่งเป็นการกลายพันธุ์บนโปรตีนหนาม และมี 3 จุดเป็นการกลายพันธุ์บนสายพันธุ์ที่น่ากังวลในเรื่องการมีผลต่อศักยภาพของการติดเชื้อและการแพร่กระจาย ได้แก่การกลายพันธุ์ N501Y ซึ่งเกิดกับ RBD ส่งผลในทางส่งเสริมการจับระหว่างโปรตีนหนามกับตัวรับ ACE2 ของคน การกลายพันธุ์ HV 69-70 deletion ส่งผลให้ไวรัสสามารถหลบหลีกระบบภูมิคุ้มกัน นอกจากนี้ การกลายพันธุ์บนตำแหน่ง P681H ซึ่งใกล้กับตำแหน่งตัดโปรตีนหนามโดยเอนไซม์ furin จึงส่งเสริมให้ไวรัสเข้ายึดเกาะและหลอมเปลือกหุ้มเข้ากับเซลล์เจ้าบ้านได้ดีขึ้น ในปัจจุบันเริ่มมีการกลายพันธุ์ E484K เกิดขึ้นกับสายพันธุ์อัลฟา เป็นที่รู้จักกันในชื่ออัลฟาพลัส¹⁹

สายพันธุ์เบต้า (B.1.351)

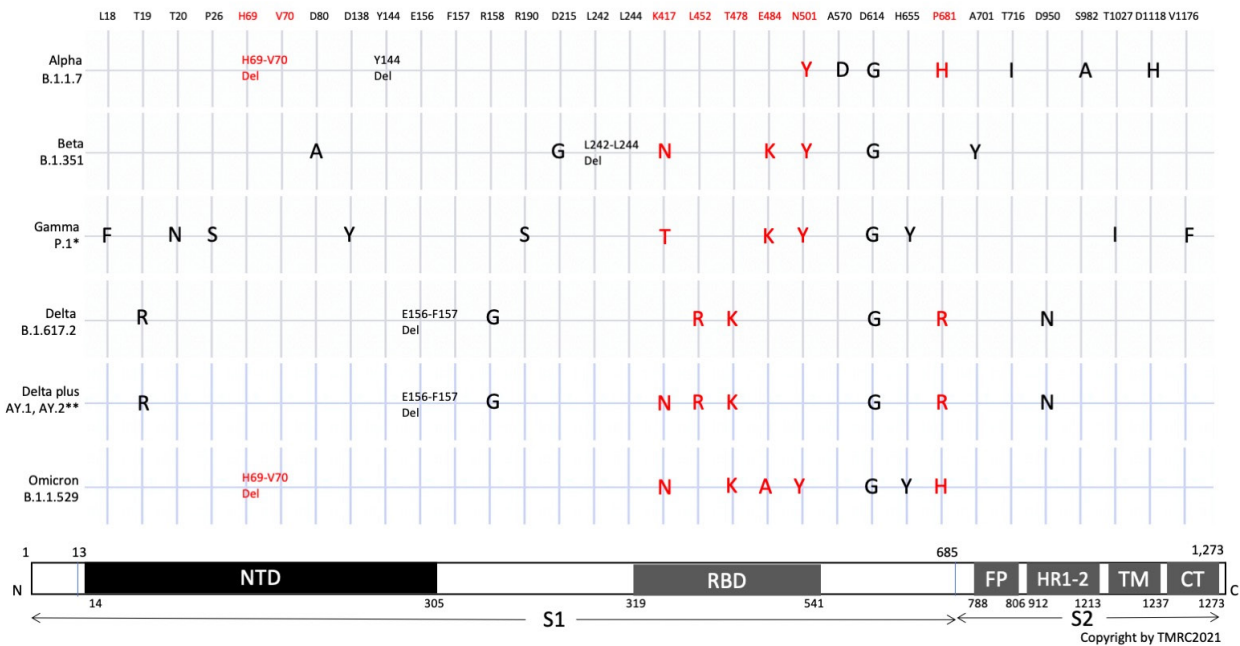
เบต้าเป็นสายพันธุ์ซึ่งเริ่มพบรายงานการระบาดเมื่อเดือนตุลาคม พ.ศ.2563 ที่ประเทศอัฟริกาใต้ และแพร่กระจายไปจนกระทั่งตรวจพบในทวีปยุโรปและสหรัฐอเมริกาในระยะถัดมา SARS-CoV-2 สายพันธุ์เบต้าประกอบด้วย การกลายพันธุ์บนโปรตีนหนาม 9 ตำแหน่ง และแม้ B.1.351 จะมีการกลายพันธุ์บน N501Y เช่นเดียวกับสายพันธุ์อัลฟา แต่สายพันธุ์ทั้งสองได้รับการพิสูจน์ทาง phylogenetics ว่ามีรากเหง้าที่แยกจากกัน และเช่นเดียวกับ N501Y การกลายพันธุ์ตำแหน่ง E484K ส่งผลให้การเข้าจับระหว่างโปรตีนหนามและตัวรับ ACE2 แน่นขึ้น²⁰ นอกจากนี้สายพันธุ์เบต้าเป็นสายพันธุ์ที่น่ากังวลเนื่องจากความสามารถในการหลบหลีกระบบภูมิคุ้มกันอันเนื่องมาจากการกลายพันธุ์บน NTD

สายพันธุ์แกมมา (P.1)

สายพันธุ์ P.1 มีต้นกำเนิดมาจาก B.1.1.28 (P มีความหมายเดียวกับ B.1.1.28) และได้รับการระบาดในกลุ่มนักท่องเที่ยวที่เดินทางจากประเทศบราซิลไปยังประเทศญี่ปุ่นและสหรัฐอเมริกา^{21,22} สายพันธุ์แกมมามีการกลายพันธุ์บนโปรตีนหนาม 10 ตำแหน่ง และ 3 ตำแหน่งในจำนวนนี้อยู่บน RBD ได้แก่ K417T E484K และ N501Y ซึ่งเป็นกลไกที่ทำให้ไวรัสเข้าสู่เซลล์ได้ง่ายขึ้น มีการติดต่อดีเร็วขึ้น และสามารถหลบหลีกภูมิคุ้มกันที่เกิดจากการได้รับวัคซีนหรือการติดเชื้อก่อนหน้า แกมมาเป็นสายพันธุ์หลักซึ่งตรวจพบในการระบาดซึ่งกลับทวีความรุนแรงขึ้นในเมือง Manaus ประเทศบราซิล ในขณะที่ประชากรมีภูมิคุ้มกันหมู่ไปแล้วประมาณร้อยละ 70 และการระบาดระลอกก่อนหน้าของบราซิลได้สงบไปแล้วราว 7 เดือน²³

สายพันธุ์เดลต้า (B.1.617.2)

สายพันธุ์เดลต้าเป็นสายพันธุ์หลักของคลื่นการระบาดที่ 3 ซึ่งเริ่มขึ้นในไตรมาสแรกของปี พ.ศ. 2564 โดยเริ่มจากประเทศอินเดียไปยังประเทศในแถบเอเชียตะวันออกเฉียงใต้รวมทั้งประเทศไทยซึ่งมีอัตราการได้รับวัคซีนต่ำในขณะนั้น นอกจากนี้ยังก่อให้เกิดการกลับมาระบาดของซ้ำในประเทศซึ่งได้รับวัคซีนไปมากแล้ว ได้แก่ อิสราเอล สหราชอาณาจักรและสหรัฐอเมริกา และเป็นเหตุให้ประเทศอิสราเอลพิจารณาเริ่มฉีดวัคซีนกระตุ้นเป็นเข็มที่สาม เนื่องจากภูมิคุ้มกันที่มีต่อเชื้อเดลต้า น้อยลง 1.4-11.3 เท่าเมื่อเทียบกับภูมิที่มีต่อ



รูปที่ 2 ตำแหน่งการกลายพันธุ์ที่เกิดกับ SARS-CoV-2 สายพันธุ์ที่น่ากังวล (Variants of Concern)

สายพันธุ์เมื่อเริ่มการระบาด การกลายพันธุ์บนโปรตีนหนามของสายพันธุ์เดลต้าประกอบด้วย L452R, T478K, D614G และ P681R โดยการกลายพันธุ์บนตำแหน่ง L452 เป็นตำแหน่งเดียวกับที่พบบนสายพันธุ์ Epsilon (B.1.429) ซึ่งได้รับการพิสูจน์ว่าแพร่พันธุ์ได้เร็วขึ้นและมีศักยภาพในการติดเชื้อสูงขึ้น 6-15 เท่า^{24,25} การกลายพันธุ์บนตำแหน่ง T478K ซึ่งอยู่บนหน้าสัมผัสระหว่างโปรตีนหนามกับตัวรับ ACE2 ส่งผลให้ไวรัสจับกับโปรตีนตัวรับได้เสถียรขึ้น นอกจากนี้สายพันธุ์เดลต้ายังมีจุดการกลายพันธุ์ P681H ซึ่งมีความสำคัญต่อการหลอมรวมเปลือกไวรัสเข้ากับผิวเซลล์และส่งเสริมการเข้าสู่เซลล์เจ้าบ้านในทางคลินิก สายพันธุ์เดลต้าเพิ่มโอกาสในการเข้ารับการรักษาในหอผู้ป่วยและอัตราตายราว 2 เท่าเมื่อเทียบกับสายพันธุ์ในกลุ่ม N501Y²⁶ ทั้งยังตอบสนองต่อวัคซีนดีกว่า²⁷

สายพันธุ์เดลต้าเริ่มมีการกลายพันธุ์ที่น่ากังวลและได้รับการจับตาในปัจจุบัน อาทิเช่น AY.1, AY.2 และ AY.3 ซึ่งมีการกลายพันธุ์บนตำแหน่ง K417N อันเป็นตำแหน่งกลายพันธุ์ที่น่ากังวลของสายพันธุ์เบต้า และ AY.4.2 ซึ่งมีตำแหน่ง A222V และ Y145H เพิ่มเติมขึ้นมา²⁸

สายพันธุ์โอไมครอน (B.1.1.529)

การระบาดของสายพันธุ์โอไมครอนเริ่มขึ้นในจังหวัดกึ่งของแอฟริกาใต้ในเดือนพฤศจิกายน 2564 จากนั้นแพร่ไปอย่างรวดเร็วจนกระทั่งมีรายงานการระบาดใน 77 ประเทศ โดยเฉพาะอย่างยิ่ง สหราชอาณาจักร แอฟริกาใต้ สหรัฐอเมริกา ภายในระยะเวลาเพียงเดือนเดียว^{1,29} การศึกษาเชิงวิวัฒนาการของไวรัสชี้ว่าโอไมครอนอาจพัฒนาสายพันธุ์มาจาก B.1.1.519 ซึ่งพบระบาดในกลุ่มผู้มีภูมิคุ้มกันบกพร่องจากการติดเชื้อไวรัส Human Immunodeficiency Virus^{2,30} นับถึงกลางเดือนธันวาคม 2564 มีข้อมูลจีโนมของสายพันธุ์โอไมครอน 4,992 รายถูกส่งไปยังฐานข้อมูล GISAID จาก 57 ประเทศ รวมถึง 3 รายจากประเทศไทย การกลายพันธุ์บนจีโนมไวรัสสายพันธุ์โอไมครอนพบมากกว่า 50 ตำแหน่งโดยประมาณ 30 ตำแหน่งเกิดขึ้นกับโปรตีนหนาม ในจำนวนนี้ 15 ตำแหน่งเกิดขึ้นกับ RBD และอาจส่งผลให้การแพร่กระจายเป็นไปอย่างรวดเร็ว การกลายพันธุ์ซึ่งพบได้ถี่ในสายพันธุ์นี้เกิดขึ้นกับตำแหน่ง H69, V70, G142, V143, Y144 และ N211 ซึ่งการขาดหายของ HV69-70 พบร่วมกับสายพันธุ์

อัลฟาและมีผลต่อการหลบหนีระบบภูมิคุ้มกันเช่นเดียวกัน การกลายพันธุ์ในบางตำแหน่งของสายพันธุ์โอไมครอนพบร่วมกับสายพันธุ์เดลต้าเช่น N501, K417 และ T478^{3,31}

การศึกษาโดยใช้แบบจำลองปัญหาประดิษฐ์ชี้ว่าการกลายพันธุ์บนตำแหน่ง N440K, T487K และ N501Y ส่งผลให้ความสามารถในการติดเชื้อของโอไมครอนสูงเป็น 10 เท่าของ SARS-CoV-2 สายพันธุ์ต้นแบบและเป็น 2 เท่าของสายพันธุ์เดลต้า นอกจากนี้ประสิทธิภาพของวัคซีนที่มีต่อโอไมครอนลดลงกว่าครึ่งเมื่อเทียบกับสายพันธุ์ที่ระบาดก่อนหน้า^{4,32} การระบาดของโอไมครอนทำให้เกิดระลอกการระบาดใหม่ในประเทศซึ่งสถานการณ์เริ่มดีขึ้นและมีการเปิดประเทศ เช่น ญี่ปุ่น อินเดีย และประเทศไทยเอง อย่างไรก็ตามแม้ปฏิบัติการติดเชื้อสายพันธุ์โอไมครอนจะสูงขึ้นมาก แต่อัตราการป่วยตายไม่สูงเท่าระลอกการระบาดเดลต้า คำอธิบายส่วนหนึ่งคือการระบาดของโอไมครอนเกิดขึ้นเมื่อประชากรได้รับวัคซีนหรือเคยติดเชื้อไปเป็นจำนวนมากแล้ว

เป้าหมายของการศึกษาลำดับเบสของสายพันธุ์ไวรัส SARS-CoV-2

การตรวจสอบลำดับสารพันธุกรรมมิได้ให้ข้อมูลซึ่งเป็นประโยชน์ต่อการรักษาโรคติดเชื้อไวรัสโคโรนา 2019 ในระดับปัจเจกบุคคล เป้าหมายของการตรวจสอบลำดับสารพันธุกรรมเป็นไปเพื่อการศึกษาเชิงระบาดวิทยา เชื่อมโยงการเปลี่ยนแปลงทางพันธุกรรมกับการเปลี่ยนแปลงของปฏิสัมพันธ์ระหว่างไวรัสและเจ้าบ้าน เพื่อเฝ้าระวังและจัดการเชิงเวชกรรมป้องกันในระดับชุมชนในกรณีที่เกิดการระบาดของสายพันธุ์ที่รุนแรง ตลอดจนสืบหาที่มาของการระบาดซึ่งในปัจจุบันยังไม่เป็นที่ทราบชัด ในปัจจุบันองค์การอนามัยโลกมีเป้าหมายในการตรวจสอบสายพันธุ์ไวรัสให้ได้ไม่น้อยกว่าร้อยละ 10 ของการระบาดทั้งหมด

การตรวจสอบสายพันธุ์ของเชื้อ SARS-CoV-2 ในประเทศไทยมีเป้าหมาย 2 ประการ ประการแรกคือสุ่มตัวอย่างเพื่อสะท้อนภาพรวมของการเปลี่ยนแปลงของสายพันธุ์ของประเทศ โดยสุ่มเก็บ 10-15 ตัวอย่างต่อเขตสุขภาพต่อสัปดาห์ ประการที่สองคือ การเฝ้าระวังสายพันธุ์ VOC/VOI โดยเก็บตัวอย่างภายใต้ข้อบ่งชี้คือ กลุ่มผู้ป่วยที่ผลการตรวจเป็นบวกและเดินทางมาจากต่างประเทศ รวมถึงต่างตัวลกลอบเข้าเมือง กลุ่มผู้ที่ได้รับวัคซีนครบแล้วยังติดเชื้อ กลุ่มผู้ที่อาการรุนแรง

และ/หรือเสียชีวิต กลุ่มผู้ติดเชื้อที่มีภาวะภูมิคุ้มกันบกพร่อง กลุ่มติดเชื้อซ้ำ กลุ่มบุคลากรทางการแพทย์ กลุ่มผู้เคยติดเชื้อมาก่อน (ติดเชื้อซ้ำ) สุ่มจากการติดเชื้อในคลัสเตอร์ใหม่ซึ่งมีจำนวนผู้ติดเชื้อตั้งแต่ 50 รายขึ้นไป หรือลักษณะทางห้องปฏิบัติการที่ชวนให้สงสัยการติดเชื้อสายพันธุ์ใหม่ เช่น การมี CT ต่ำกว่าปกติ

การตรวจคัดกรองสายพันธุ์ของเชื้อ SARS CoV-2

การตรวจพบสายพันธุ์ใหม่ๆ ของเชื้อไวรัส SARS CoV-2 ที่เกิดจากปรับเปลี่ยนตัวเองตลอดเวลา นับเป็นปัจจัยสำคัญที่ต้องคำนึงถึงในการวางแผนป้องกันการแพร่ระบาดของโรคติดเชื้อไวรัสโคโรนา 2019 หากทำการตรวจพบสายพันธุ์ใหม่ได้รวดเร็วจะทำให้สามารถควบคุมโรคได้ดียิ่งขึ้น ปัจจุบันเทคโนโลยีที่ได้รับการยอมรับเป็นการทดสอบมาตรฐานในการตรวจหาสายพันธุ์ในปัจจุบันคือการถอดรหัสสารพันธุกรรม Next Generation Sequencing (NGS) ซึ่งเป็นตรวจสอบสายพันธุกรรมที่ช่วยให้สามารถพบการเปลี่ยนแปลงของเชื้อได้อย่างละเอียด แต่จำเป็นต้องใช้วัสดุอุปกรณ์ที่มีราคาสูง และต้องใช้ความชำนาญในการวิเคราะห์เป็นอย่างมาก จึงต้องใช้ต้นทุนสูงและระยะเวลานาน ทำให้เริ่มมีการพัฒนาเทคนิคอื่น ๆ ที่สามารถใช้งานได้สะดวกขึ้นมาใช้ในการตรวจคัดกรองเบื้องต้น เช่น

การตรวจด้วยวิธี real-time reverse transcriptase Polymerase Chain Reaction-methods (RT-PCR) เป็นเทคนิคการเพิ่มปริมาณสารพันธุกรรมเป้าหมาย โดยการใช้การติดตามวัดสัญญาณเรืองแสงที่เกิดขึ้นเมื่อสารพันธุกรรมเป้าหมาย (PCR product) เพิ่มจำนวน เพื่อหาตำแหน่งการกลายพันธุ์ที่ได้มีการระบุแล้วว่าตำแหน่งนั้นสามารถพบในสายพันธุ์ใด โดยเริ่มจากขั้นตอนการเปลี่ยน RNA ให้เป็น DNA (Reverse transcription) และเข้าสู่ขั้นตอนการเพิ่มปริมาณสารพันธุกรรม โดยอาศัยปฏิกิริยาลูกโซ่พอลิเมอเรส (PCR) ด้วยชุดไพรเมอร์ที่ประกอบด้วย forward primers, reverse primers ของตำแหน่งเป้าหมาย (target DNAs) และ Probe ซึ่งเป็น DNA สายสั้นที่มีการติดสารเรืองแสง (reporter) ไว้ที่ปลาย 5' ที่มีสารบ่งแสง (Quencher) ติดที่ปลาย 3' เมื่อมีการสร้าง DNA สายใหม่ primer จะไปชน probe ที่จับอยู่บนเส้น DNA เดียวกัน probe จะถูกย่อยทำให้ reporter จะออกมาอยู่ห่างจาก quencher จึงตรวจพบสัญญาณ

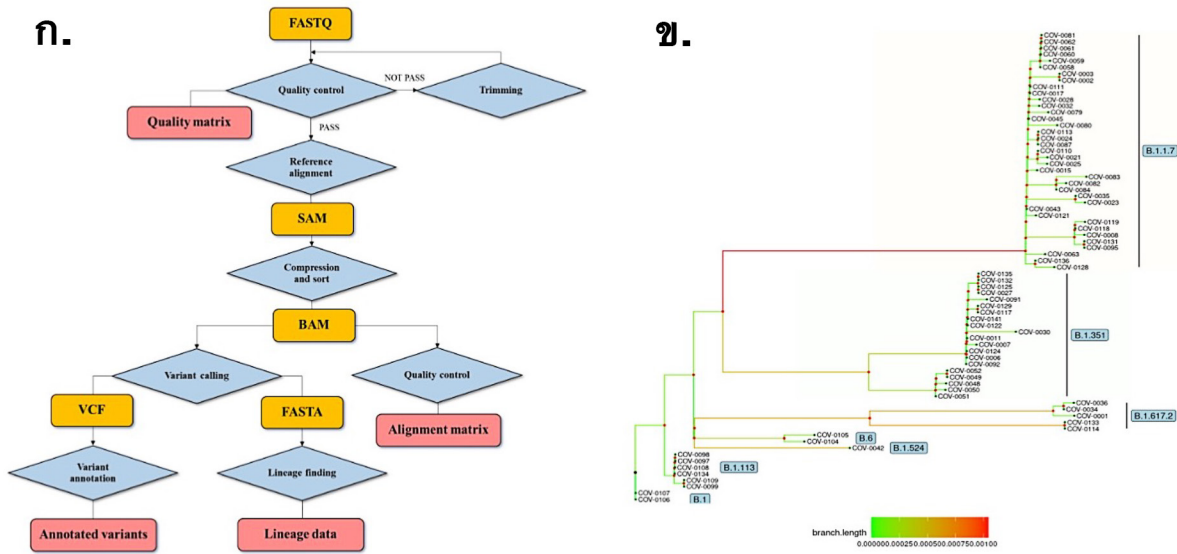
เรืองแสงได้ โดยสามารถตรวจได้หลายเป้าหมายพร้อมกันโดยการเลือก reporter ที่ให้การเรืองแสงในช่วงค่าความถี่ต่างกัน ปัจจุบันมีชุดตรวจทางการค้าออกมาจำหน่ายไม่น้อยโดยเน้นการตรวจวิเคราะห์ตำแหน่งการกลายพันธุ์ในยีนที่กำหนดการสร้างโครงสร้างหนาม (spike protein) ที่มีบทบาทสำคัญในการจับกับเซลล์เจ้าบ้าน โดยตำแหน่งที่มักจะมีการตรวจอย่างเช่น L452R, E484Q, N501Y, E484K, P681R, K417T, K417N และ HV69-70del ซึ่งในการระบุสายพันธุ์ให้มีความแม่นยำจำเป็นต้องทราบตำแหน่งการกลายพันธุ์ที่มากกว่า 1 ตำแหน่งมาพิจารณาาร่วมกันเช่น การพบการกลายพันธุ์ที่ L452R และ P681K เป็นสองตำแหน่งที่พบได้ในสายพันธุ์ Delta ในขณะที่หากมีการพบ N501Y และ HV69-70 del ซึ่งทั้ง 2 ตำแหน่งสามารถพบได้ทั้งใน Alpha และ Omicron ทำให้ต้องมีการตรวจหาตำแหน่งการกลายพันธุ์หลายตำแหน่งมากขึ้นที่มีความจำเพาะและแตกต่าง ซึ่งเป็นข้อจำกัดอย่างหนึ่งของวิธี Real Time RT-PCR ที่ต้องหา reporter ที่เรืองแสงในช่วงค่าความถี่ต่างกันให้พอกับจำนวนตำแหน่งการกลายพันธุ์

การตรวจด้วยเทคโนโลยี Mass Array จึงเป็นอีกเทคนิคหนึ่งที่ได้รับการพัฒนาใช้สำหรับการตรวจวิเคราะห์สายพันธุ์กรรม ซึ่งเป็นการวิเคราะห์ที่ประกอบด้วย 2 หลักการคือ Multiplex PCR ร่วมกับ MALDI-TOF Mass Spectrometry ในขั้นตอนการทำ Multiplex PCR ประกอบด้วย 3 ขั้นตอน ได้แก่ PCR Amplification คือ ขั้นตอนการเพิ่มปริมาณสารพันธุกรรม โดยอาศัยปฏิกิริยาถูกโซ่พอลิเมอเรส (PCR) ด้วยชุดไพรเมอร์ที่ประกอบด้วย forward primers และ reverse primers ของตำแหน่งเป้าหมาย (target DNAs) ตามด้วย Shrimp Alkaline Phosphatase (SAP) Treatment คือ ขั้นตอนการกำจัด dNTPs ส่วนเกินที่เหลือจากการทำ PCR Amplification ในขั้นตอนแรก และ Single Base Extension (SBE) คือ ขั้นตอนปฏิกิริยาการต่อนิวคลีโอไทด์เบสที่มีคุณสมบัติเป็น terminator nucleotides (ddNTPs) จำนวน 1 เบสเข้าไปยังฝั่ง 3' ของสาย extension primers ที่ออกแบบมาจำเพาะให้อยู่ติดกับตำแหน่งเป้าหมาย เพื่อให้ terminator nucleotides ซึ่งมีคุณลักษณะเป็นเบสคู่สมกับตำแหน่งเป้าหมายมาต่อ จากนั้นจึงหยุดการต่อสายทันทีเกิดผลิตภัณฑ์ขึ้นเรียกว่า extension products หลังจากนั้นจึงนำ extension products เข้าสู่กระบวนการตรวจวิเคราะห์

ด้วย Mass Spectrometry ต่อไป การทำ MALDI-TOF Mass Spectrometry Analysis เป็นการตรวจวิเคราะห์วัดน้ำหนักมวลโมเลกุลของ extension products ซึ่งมีน้ำหนักมวลโมเลกุลแตกต่างกันขึ้นอยู่กับชนิดของสารพันธุกรรมและ nucleotide base ตรงตำแหน่ง variant ที่เกิดขึ้น สารที่น้ำหนักน้อยจะวิ่งได้เร็วกว่าสารที่มีน้ำหนักมาก ทำให้แยกสารตัวอย่างที่ต้องการหาออกมาเป็น Peak ที่เห็นได้ชัดเจน โดยทำการวิเคราะห์พร้อมประมวลผลด้วยซอฟต์แวร์ การใช้ MassARRAY ในการวิเคราะห์สารที่รู้ลำดับเบสแน่นอน (Known Target) จึงมีความแม่นยำถูกต้องในระดับร้อยละ 99.7 และยังสามารถตรวจการกลายพันธุ์ได้หลายตำแหน่ง ซึ่งปัจจุบันได้มีการพัฒนาให้มีการตรวจได้กว่า 30 ตำแหน่งที่มีความจำเพาะเพื่อใช้ในการวิเคราะห์สายพันธุ์ ซึ่งให้ความเชื่อมั่นในการประมวลผลการคัดกรองสายพันธุ์ได้แม่นยำยิ่งขึ้น

ขั้นตอนการวิเคราะห์ทางชีวสารสนเทศในการตรวจสอบลำดับสารพันธุกรรมไวรัส SARS-CoV-2

การวิเคราะห์จีโนมของเชื้อไวรัสโคโรนาสายพันธุ์ใหม่ มีวัตถุประสงค์เพื่อการลดทอนขนาดข้อมูลรหัสพันธุกรรมจากกระบวนการถอดรหัสพันธุกรรม ให้เป็นข้อมูลรหัสพันธุกรรมสายเดี่ยว (consensus FASTA) และข้อมูลตำแหน่งการกลายพันธุ์บนจีโนมของไวรัส (variant calling format-VCF) ที่เพียงพอสำหรับระบุสายพันธุ์ของเชื้อไวรัสที่ทราบรูปแบบการกลายพันธุ์แล้ว และการเฝ้าระวังการกลายพันธุ์ใหม่ ในเบื้องต้นภายหลังการถอดรหัสพันธุกรรม จะได้รับข้อมูลรหัสพันธุกรรม (nucleotide base) ควบคู่กับข้อมูลคุณภาพ (base quality score) ของแต่ละชุดข้อมูล (reads) อิงตามรูปแบบการเตรียมสารพันธุกรรม (library preparation) และมีขนาดความยาวตามการตัดแบ่งสายพันธุกรรม โดยข้อมูลทั้งหมดถูกจัดเก็บในรูปแบบไฟล์ FASTQ ข้อมูลรหัสพันธุกรรมดังกล่าวควรได้รับการตรวจสอบคุณภาพ (quality control) ก่อนดำเนินการวิเคราะห์ในขั้นต่อไป (รูปที่ 3 ก.) โดยกระบวนการกรองคุณภาพข้อมูลประกอบด้วย¹ คุณภาพของรหัสพันธุกรรมแต่ลำดับโดยภาพรวม² ขนาดความยาวของรหัสพันธุกรรม (read length)³ สัดส่วนของนิวคลีโอไทด์และการกระจายตัวของคู่เบส G-C (guanine-cytosine)⁴ ปริมาณข้อมูลที่ถูกถอดรหัสซ้ำ (sequence duplication)⁵ และปริมาณข้อมูลนิวคลีโอไทด์สังเคราะห์ที่ใช้เพื่อการติดตามและเชื่อมต่อกัน



SAM = sequence alignment map; BAM = binary alignment map; VCF = variant calling format

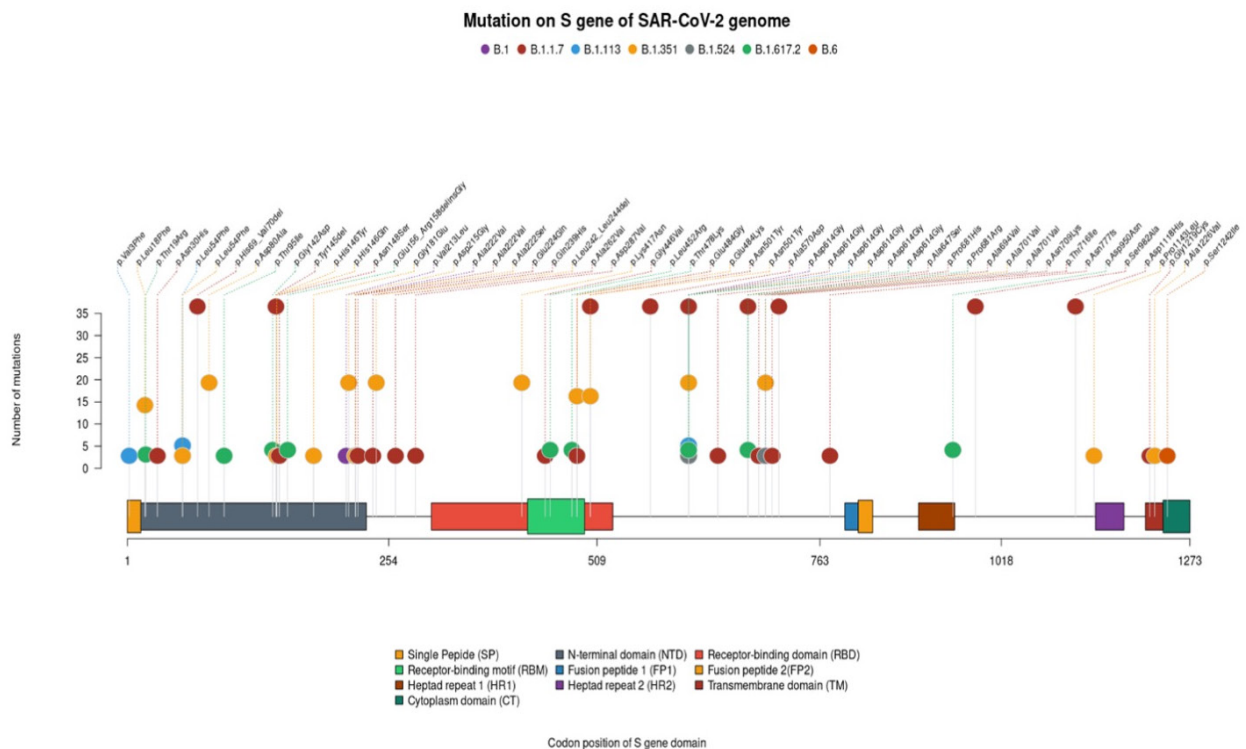
รูปที่ 3 ก. ผังงานทางชีวสารสนเทศในการวิเคราะห์ข้อมูลทางพันธุกรรมของไวรัส SARS-CoV-2

ข. ตัวอย่างแผนภูมิแสดงต้นไม้วิวัฒนาการที่ได้จากการตรวจสอบลำดับสารพันธุกรรมของการระบาดในภาคใต้ในช่วงเดือนตุลาคม พ.ศ. 2564

ยังหลงเหลือ (remaining sequence adapter and index) หากข้อมูลรหัสพันธุกรรมบางส่วนผ่านเกณฑ์มาตรฐาน กล่าวคือ ควรมีค่าคุณภาพโดยภาพรวมสูงกว่า Q30 ควรมีปริมาณข้อมูลที่ถูกต้องรหัสอย่างน้อยกว่า 20 ขั้วต่อชุดข้อมูล และตรวจไม่พบข้อมูลนิวคลีโอไทด์สังเคราะห์ จะทำให้การวิเคราะห์มีความคลาดเคลื่อนน้อยลงหรือมีความน่าเชื่อถือเพิ่มขึ้น อย่างไรก็ตามข้อมูลรหัสพันธุกรรมที่ไม่ผ่านเกณฑ์มาตรฐาน อาจปรับแก้คุณภาพได้บางส่วนโดยการตัดเล็มข้อมูลที่มีคุณภาพต่ำทิ้ง รวมถึงตัดข้อมูลนิวคลีโอไทด์สังเคราะห์ และลดรูปชุดข้อมูลที่ถูกต้องรหัสซ้ำให้เหลือเพียงชุดข้อมูลเดียว เมื่อข้อมูลรหัสพันธุกรรมผ่านเกณฑ์มาตรฐานแล้วจึงสามารถนำไปเรียงลำดับเปรียบเทียบกับจีโนมไวรัสอ้างอิง ซึ่งสำหรับจีโนมอ้างอิงของไวรัส SARS-CoV-2 เป็นข้อมูลที่ได้รับจากการถอดรหัสพันธุกรรมสายยาว (long read sequencing) ของสายพันธุ์ดั้งเดิม

ข้อมูลรหัสพันธุกรรมที่ถูกจัดวางบนจีโนมอ้างอิงจะบันทึกในรูปแบบไฟล์ SAM (sequence alignment map) และเพื่อให้ประมวลผลข้อมูลทำได้อย่างรวดเร็วผนวกกับประหยัด

พื้นที่จัดเก็บ การลดทอนขนาดไฟล์และการเรียงลำดับจึงเข้ามา มีบทบาทสำคัญภายหลังจากการจัดเรียง ข้อมูลจะถูกแปลงให้อยู่ในรูปแบบ BAM (binary alignment map) ลำดับถัดไปข้อมูลรหัสพันธุกรรมจะถูกนำไปวิเคราะห์เพื่อค้นหาตำแหน่งการกลายพันธุ์ซึ่งเป็นกระบวนการตรวจสอบลำดับเบสที่แตกต่างระหว่างข้อมูลรหัสพันธุกรรมที่วิเคราะห์กับจีโนมอ้างอิง นอกจากนี้ข้อมูลตำแหน่งการกลายพันธุ์ยังเป็นข้อมูลพื้นฐานสำหรับสร้างข้อมูลรหัสพันธุกรรมสายเดี่ยวสำหรับใช้สร้างแผนภูมิต้นไม้ (phylogenetic tree) แสดงในรูปที่ 3 ข. เพื่อระบุสายพันธุ์ของเชื้อไวรัส ซึ่งอิงตำแหน่งการกลายพันธุ์ที่มีความน่าเชื่อถือสูงควรมีคุณภาพของลำดับเบสเฉลี่ยไม่น้อยกว่า Q30 มีชุดข้อมูลรหัสพันธุกรรมสนับสนุน (reading depth) ไม่น้อยกว่า 20 ชุด และควรมีความถี่อัลลีลที่กลายพันธุ์ไม่น้อยกว่าร้อยละ 90 นอกจากนี้ข้อมูลตำแหน่งการกลายพันธุ์ยังสามารถนำมาวิเคราะห์เพิ่มเติมโดยเป็นการระบุผลกระทบของการกลายพันธุ์ต่อการเปลี่ยนแปลงลำดับกรดอะมิโน (variant annotation) สามารถนำมาแสดงผลในเชิงของกราฟการกลายพันธุ์ (mutation landscape) รูปที่ 4 ข้อมูล



รูปที่ 4 แสดงตัวอย่างตำแหน่งการกลายพันธุ์ที่ตรวจพบในตัวอย่างไวรัส SARS-CoV-2 ที่ระบาดในภาคใต้ในช่วงเดือนตุลาคม 2564

คำจำกัดความของสายพันธุ์ที่ได้รับการ update สามารถติดตามได้จากฐานข้อมูลขององค์การอนามัยโลก <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/>

สรุป

การระบาดของโรคติดเชื้อไวรัสโคโรนา 2019 ซึ่งเริ่มต้นขึ้นในปี พ.ศ.2563 และยังไม่มียุติสิ้นสุดในขณะนี้มิได้เกิดจากไวรัส SARS-CoV-2 ตัวเดิมซึ่งเริ่มการระบาดขึ้นในเมืองอู่ฮั่น ในขณะที่คลื่นของการระบาดเคลื่อนออกไปตามมิติทางภูมิศาสตร์และเวลา วิวัฒนาการของสายพันธุ์เกิดขึ้นและเป็นไปในทิศทางซึ่งทำให้ไวรัสมีศักยภาพสูงขึ้นในการติดเชื้อ หลบหลีกระบบภูมิคุ้มกันและการไล่ล่าของวัคซีน การตรวจสอบสายพันธุ์ของไวรัสซึ่งจะเข้าใจ ติดตามการกลายพันธุ์ของไวรัส และสร้างภูมิคุ้มกันให้ทันกับการปรับตัวดังกล่าว

กิติกรรมประกาศ

การเรียบเรียงบทความนี้ได้รับการสนับสนุนจากโครงการความร่วมมือในการตรวจหาสายพันธุ์ไวรัส SARS-CoV-2 ระหว่างมหาวิทยาลัยสงขลานครินทร์และกรมวิทยาศาสตร์การแพทย์ ผ่านศูนย์ตรวจวิเคราะห์สายพันธุ์ไวรัสโคโรนาสายพันธุ์ใหม่ภาคใต้ (Center for Emerging SARS-CoV-2 Lineage Investigation in Southern Thailand, CESLIST)

เอกสารอ้างอิง

1. Jamison DT, Wu KB. The East-West Divide in Response to COVID-19. Engineering (Beijing) 2021;7:936-47.
2. Calleja N, AbdAllah A, Abad N, Ahmed N, Albarracin D, Altieri E, et al. A Public Health Research Agenda for Managing Infodemics: Methods and Results of the First WHO Infodemiology Conference. JMIR Infodemiology 2021;1:e30979.

3. Kalia K, Saberwal G, Sharma G. The lag in SARS-CoV-2 genome submissions to GISAID. *Nat Biotechnol* 2021;39:1058–60.
4. Van Noorden R. Scientists call for fully open sharing of coronavirus genome data. *Nature* 2021;590:195–6.
5. Paul P, France AM, Aoki Y, Batra D, Biggerstaff M, Dugan V, et al. Genomic Surveillance for SARS-CoV-2 Variants Circulating in the United States, December 2020–May 2021. *MMWR Morb Mortal Wkly Rep* 2021;70:846–50.
6. Wang H, Li X, Li T, Zhang S, Wang L, Wu X, et al. The genetic sequence, origin, and diagnosis of SARS-CoV-2. *Eur J Clin Microbiol Infect Dis* 2020;39:1629–35.
7. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. *Nature* 2020;579:265–9.
8. who.int [homepage of the Internet]. Geneva: WHO; Weekly operational update on COVID-19 – 25 October 2021 [cited 2021 October 25]. Available from: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>.
9. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol* 2021;19:409–24.
10. Boopathi S, Poma AB, Kolandaivel P. Novel 2019 coronavirus structure, mechanism of action, antiviral drug promises and rule out against its treatment. *J Biomol Struct Dyn* 2021;39:3409–18.
11. Manzaneres-Meza LD, Medina-Contreras O. SARS-CoV-2 and influenza: a comparative overview and treatment implications. *Bol Med Hosp Infant Mex* 2020;77:262–73.
12. Jogalekar MP, Veerabathini A, Gangadaran P. Novel 2019 coronavirus: Genome structure, clinical trials, and outstanding questions. *Exp Biol Med (Maywood)* 2020;245:964–9.
13. Huang Y, Yang C, Xu XF, Xu W, Liu SW. Structural and functional properties of SARS-CoV-2 spike protein: potential antiviral drug development for COVID-19. *Acta Pharmacol Sin* 2020;41:1141–9.
14. Wang MY, Zhao R, Gao LJ, Gao XF, Wang DP, Cao JM. SARS-CoV-2: Structure, Biology, and Structure-Based Therapeutics Development. *Front Cell Infect Microbiol* 2020;10:587269.
15. Letko M, Marzi A, Munster V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol* 2020;5:562–9.
16. Heurich A, Hofmann-Winkler H, Gierer S, Liepold T, Jahn O, Pohlmann S. TMPRSS2 and ADAM17 cleave ACE2 differentially and only proteolysis by TMPRSS2 augments entry driven by the severe acute respiratory syndrome coronavirus spike protein. *J Virol* 2014;88:1293–307.
17. Rambaut A, Holmes EC, O’Toole A, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol* 2020;5:1403–7.
18. Shu Y, McCauley J. GISAID: Global initiative on sharing all influenza data – from vision to reality. *Euro Surveill* 2017;22:30494.
19. covariants.org [homepage of the Internet]. Bern: Emma Hodcroft; Variant:20I (Alpha, V1) [cited 27 October 2021]. Available from: <https://covariants.org/variants/20I.Alpha.V1>.
20. Starr TN, Greaney AJ, Hilton SK, Ellis D, Crawford KHD, Dingens AS, et al. Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ace2 binding. *Cell* 2020;182:1295–310.
21. Firestone MJ, Lorentz AJ, Meyer S, Wang X, Como-Sabetti K, Vetter S, et al. First Identified Cases of SARS-CoV-2 Variant P.1 in the United States – Minnesota, January 2021. *MMWR Morb Mortal Wkly Rep* 2021;70:346–7.
22. Fujino T, Nomoto H, Kutsuna S, Ujiie M, Suzuki T, Sato R, et al. Novel SARS-CoV-2 Variant in Travelers from Brazil to Japan. *Emerg Infect Dis* 2021;27.
23. Sabino EC, Buss LF, Carvalho MPS, Prete CA, Jr, Crispim MAE, Fraiji NA, et al. Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. *Lancet* 2021;397:452–5.
24. Deng X, Garcia-Knight MA, Khalid MM, Servellita V, Wang C, Morris MK, et al. Transmission, infectivity, and antibody neutralization of an emerging SARS-CoV-2 variant in California carrying a L452R spike protein mutation. *medRxiv*. 2021.
25. Motozono C, Toyoda M, Zahradnik J, Saito A, Nasser H, Tan TS, et al. SARS-CoV-2 spike L452R variant evades cellular immunity and increases infectivity. *Cell Host Microbe* 2021;29:1124–36.
26. Fisman DN, Tuite AR. Evaluation of the relative virulence of novel SARS-CoV-2 variants: a retrospective cohort study in Ontario, Canada. *CMAJ* 2021;193:E1619–25.
27. Liu C, Ginn HM, Dejnirattisai W, Supasa P, Wang B, Tuekprakhon A, et al. Reduced neutralization of SARS-CoV-2 B.1.617 by vaccine and convalescent serum. *Cell* 2021;184:4220–36.
28. covariants.org [homepage of the Internet]. Bern: Emma Hodcroft; Variant: 21A (Delta) [cited 2021 October]. Available from: <https://covariants.org/variants/21A.Delta>.

29. Al-Tawfiq JA, Hoang VT, Le Bui N, Chu DT, Memish ZA. The Emergence of the Omicron (B.1.1.529) SARS-CoV-2 Variant: What is the Impact on the Continued Pandemic? *J Epidemiol Glob Health* 2022;1-4.
30. Callaway E. Heavily mutated Omicron variant puts scientists on alert. *Nature* 2021;600:21.
31. Chen J, Wang R, Gilby NB, Wei GW. Omicron Variant (B.1.1.529): Infectivity, Vaccine Breakthrough, and Antibody Resistance. *J Chem Inf Model* 2022;62:412-22.
32. Ai J, Zhang H, Zhang Y, Lin K, Zhang Y, Wu J, et al. Omicron variant showed lower neutralizing sensitivity than other SARS-CoV-2 variants to immune sera elicited by vaccines after boost. *Emerg Microbes Infect* 2022;11:337-43.