



แนวทางการใช้ชีวสารสนเทศกับผลตรวจดีเอ็นเอในงานนิติวิทยาศาสตร์ Guidelines for Using Bioinformatics in Forensic DNA Profiles

พรวิภา เวฬุกัญญา และ สุพรรณ มลิกหา
คณะนิติวิทยาศาสตร์ โรงเรียนนายร้อยตำรวจ

Phornvipa Werukanjana and Supunsa Malikhao
Faculty of Forensic Science, Royal Police Cadet Academy

Received July 27, 2021 | Revised October 1, 2021 | Accepted October 27, 2021

บทความวิชาการ (Academic Article)

บทคัดย่อ

ปัจจุบันเทคโนโลยีชีวสารสนเทศของดีเอ็นเอในงานนิติวิทยาศาสตร์ถูกนำมาใช้เพื่อช่วยในการสอบสวนคดีอาญา ข้อดีของเทคนิคนี้ คือ สามารถเชื่อมโยงเพื่อตรวจสอบและระบุตัวบุคคลที่มีประสิทธิภาพสูง ปัจจุบันมีหน่วยงานที่รวบรวมจัดเก็บชีวสารสนเทศของดีเอ็นเอสำหรับอ้างอิงในการวิจัยและงานด้านนิติวิทยาศาสตร์ตามมาตรฐานของสถาบันมาตรฐานและเทคโนโลยีแห่งชาติสหรัฐอเมริกา (NIST) กำหนด โดยชีวสารสนเทศของดีเอ็นเอสามารถนำมาใช้เพื่อการวิจัย การทดลองและการทดสอบโดยใช้เป็นข้อมูลตัวอย่างในการสร้างระบบฐานข้อมูลดีเอ็นเอของตนเองก่อนใช้กับข้อมูลจริงเพื่ออำนวยความสะดวกในการทำงานในห้องปฏิบัติการสำหรับองค์กรที่ไม่เป็นสมาชิกของระบบดัชนีดีเอ็นเอ (CODIS) ปัจจุบันหน่วยงานที่ใช้ฐานข้อมูลดีเอ็นเอสำหรับการค้นหาเปรียบเทียบดีเอ็นเอระหว่างประเทศ เช่น องค์การตำรวจอาชญากรรมระหว่างประเทศ (INTERPOL) สำนักงานสอบสวนกลางแห่งสหรัฐอเมริกา (FBI) แต่ยังมีหน่วยงานอีกหลายหน่วยงานที่ไม่ได้เป็นสมาชิกระบบดัชนีดีเอ็นเอซึ่งจะไม่สามารถเข้าใช้ระบบดังกล่าวได้ จึงเป็นเหตุให้หน่วยงานที่ไม่ได้เป็นสมาชิกระบบดัชนีดีเอ็นเอจะต้องสร้างระบบฐานข้อมูลเพื่อทำการเปรียบเทียบดีเอ็นเอขึ้นมาเอง และเพื่อไม่ให้เกิดผลกระทบกับข้อมูลสำคัญที่มีอยู่เดิมจึงจำเป็นต้องหาแหล่งข้อมูลดีเอ็นเอมาใช้เพื่อสร้างระบบฐานข้อมูลดีเอ็นเออย่างถูกต้องและมีรูปแบบตรงกับระบบการตรวจพิสูจน์ที่มีอยู่อย่างมีมาตรฐาน บทความนี้จะกล่าวถึงองค์กรที่เกี่ยวข้อง เช่น องค์กรที่กำหนดมาตรฐานที่ใช้อ้างอิงแหล่งข้อมูลดีเอ็นเอที่นำมาใช้เพื่อการวิจัยและทดสอบเบื้องต้นสำหรับการเปรียบเทียบดีเอ็นเอ ตลอดจนแนะแนวทางในการสร้างฐานข้อมูลชีวสารสนเทศของดีเอ็นเอสำหรับหน่วยงานที่ไม่ได้เป็นสมาชิกระบบดัชนีดีเอ็นเอ ตั้งแต่หาแหล่งข้อมูลดีเอ็นเอก่อนที่จะนำไปออกแบบฐานข้อมูลกับชีวสารสนเทศของดีเอ็นเอโดยไม่กระทบกับข้อมูลดีเอ็นเอที่มีอยู่เดิม โดยจะกล่าวถึงพรรณนาที่ไม่ลงลึกทางเทคนิคมากนักตั้งแต่เริ่มต้นจนกระทั่งได้แนวทางไปสู่การสร้างฐานข้อมูลชีวสารสนเทศของดีเอ็นเอของหน่วยงานที่เป็นไปตามมาตรฐาน

คำสำคัญ: ชีวสารสนเทศ, ผลตรวจดีเอ็นเอ, ฐานข้อมูลดีเอ็นเอ, ระบบดัชนีดีเอ็นเอ

Abstract

Nowadays, Bioinformatics is applied for the analyses of forensic DNA profile that has become one of the most foremost and significant tools of crime investigations. Bioinformatics have revolutionized the ability to use DNA profile information to link crime scene evidence to perpetrators. The advantages of this technique can be linked to highly effective authentication and identification. According to the standards of the United States National Institute of Standards and Technology (NIST) that DNA bioinformatics can be used for research and forensic science. Trials and tests are used as a sample data to create their own DNA database system before entering the actual data. Currently, some agencies use DNA databases for international DNA comparison searches, such as the International Criminal Police Organization (INTERPOL) and the Federal Bureau of Investigation (FBI) but many agencies are not members they will not be able to access the system. Therefore, agencies that are not members of the DNA index system will create their own database to use and compare DNA in order not to affect the existing important information. It is necessary to find a source of DNA to create a DNA bioinformatics database system that is accurate and consistent with existing standard verification systems. This article discusses the organizations involved such as organizations that provide reference standards DNA resources used for preliminary research and testing for DNA comparisons. As well as guidance on creating a bioinformatics database of DNA for the organizations are not members of the DNA index system. DNA Database does not deep into technical depth by presenting guidelines for database creation DNA bioinformatics for building a self-organized DNA bioinformatics database.

Keywords: Bioinformatics, DNA Profile, DNA Database, CODIS

บทนำ

บทความนี้จะกล่าวถึงแนวทางในการใช้ชีวสารสนเทศกับผลการตรวจดีเอ็นเอในงานนิติวิทยาศาสตร์ การตรวจดีเอ็นเอนั้นมีคุณูปการต่องานนิติวิทยาศาสตร์และกระบวนการยุติธรรม โดยผลการตรวจดีเอ็นเอสามารถจัดเก็บไว้ในฐานข้อมูลในรูปแบบชีวสารสนเทศของดีเอ็นเอ ซึ่งแต่ละประเทศสามารถใช้รูปแบบข้อมูลที่สามารถแบ่งปันข้อมูลกันระหว่างหน่วยงานที่เป็นสมาชิกภายใต้เงื่อนไขการใช้งานของระบบได้ เช่น สำนักงานสอบสวนกลางสหรัฐอเมริกา (Federal Bureau of Investigation; FBI) (FBI, 2008) และองค์การตำรวจอาชญากรรมระหว่างประเทศ (International Criminal Police Organization; INTERPOL) หรือตำรวจสากล (INTERPOL, 2019) ทั้งสององค์กรนี้ใช้ระบบดัชนีดีเอ็นเอ (Combine DNA Index System; CODIS) ในการค้นหาและแลกเปลี่ยนข้อมูลระหว่างชาติสมาชิก

เนื้อหาส่วนที่ 1 กล่าวถึงการตรวจดีเอ็นเอในงานนิติวิทยาศาสตร์ ตัวอย่างข้อมูลดีเอ็นเอ การใช้เทคโนโลยีที่สนับสนุนการวิเคราะห์และจัดเก็บดีเอ็นเอ ซึ่งผลการตรวจดีเอ็นเอที่ได้จะอยู่ในรูปชีวสารสนเทศชนิดเครื่องหมายดีเอ็นเอ STR (Short tandem repeat) ผลการตรวจยืนยันใน



ห้องปฏิบัติการประกอบด้วยข้อมูลสำคัญในการระบุตัวบุคคล การค้นหาและเปรียบเทียบดีเอ็นเอของผู้กระทำความผิดกับหลักฐานในที่เกิดเหตุ เพื่อช่วยให้การสืบสวนสอบสวนทำงานอย่างมีประสิทธิภาพ เพราะการตรวจดีเอ็นเอสามารถยืนยันตัวตนได้อย่างแม่นยำ

เนื้อหาส่วนที่ 2 กล่าวถึงหน่วยงานที่เป็นแหล่งข้อมูลด้านชีวสารสนเทศของดีเอ็นเอ หน่วยงานต่าง ๆ ที่เป็นแหล่งชีวสารสนเทศที่กระจายอยู่ทั่วโลกและหน่วยงานชีวสารสนเทศในประเทศไทย หน่วยงานดังกล่าวนี้ตั้งขึ้นมาเพื่อการวิจัยและเป็นแหล่งอ้างอิง มุ่งเน้นด้านชีวสารสนเทศรูปแบบต่าง ๆ ในบทความนี้จะเน้นแหล่งข้อมูลด้านชีวสารสนเทศของดีเอ็นเอที่เกี่ยวข้องกับงานด้านนิติวิทยาศาสตร์ เพื่อเป็นแนวทางให้เห็นถึงความสำคัญที่แต่ละองค์กรทั่วโลกได้สร้างขึ้น

เนื้อหาส่วนที่ 3 กล่าวถึงหน่วยงานกำหนดมาตรฐานข้อมูลชีวสารสนเทศดีเอ็นเอในงานนิติวิทยาศาสตร์ที่เกี่ยวข้องกับแนวทางในการนำเทคโนโลยีพื้นฐานโดยอ้างอิงสถาบันมาตรฐานและเทคโนโลยีแห่งชาติสหรัฐอเมริกา (National Institute of Standards and Technology; NIST) และหน่วยงานที่ใช้ระบบมาตรฐานสำหรับข้อมูลดีเอ็นเอ เช่น ระบบดัชนีดีเอ็นเอ (CODIS) ซึ่งมีหลายองค์กรและหลายประเทศเป็นสมาชิก แต่หลายองค์กรและหลายประเทศไม่ได้เลือกใช้ระบบนี้ เพราะแต่ละประเทศหรือองค์กร อาจมีกฎหมายที่ไม่อนุญาตให้เปิดเผยข้อมูลส่วนบุคคลด้วยเหตุผลทางจริยธรรม องค์กรต่าง ๆ เหล่านี้จะต้องสร้างระบบฐานข้อมูลชีวสารสนเทศของดีเอ็นเอเพื่อการใช้งานขึ้นมาเองและควรอ้างอิงมาตรฐานจากหน่วยงานมาตรฐานดังกล่าวมา

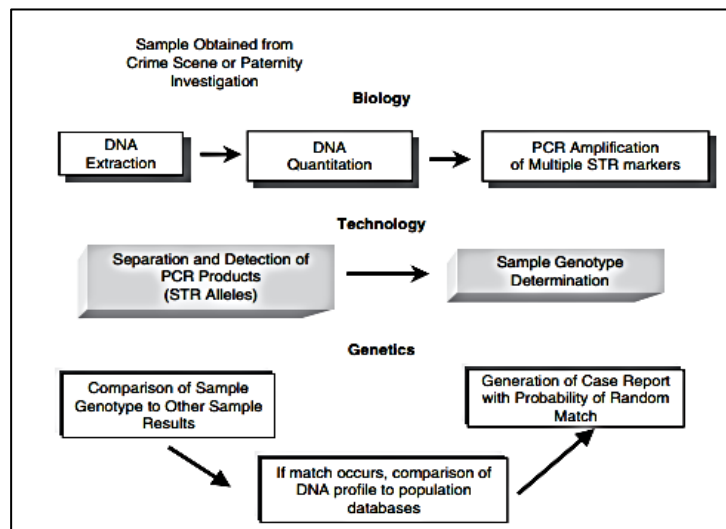
เนื้อหาส่วนที่ 4 กล่าวถึงหน่วยงานที่ใช้ประโยชน์จากฐานข้อมูลชีวสารสนเทศของดีเอ็นเอ ซึ่งมีบทบาทสำคัญในการนำไปแก้ปัญหาอาชญากรรม เนื่องจากมีศักยภาพในการเชื่อมโยงกลุ่มอาชญากรรมและ/หรือ กำหนดตัวผู้ต้องสงสัยในที่เกิดเหตุโดยใช้วิธีการทางนิติวิทยาศาสตร์ ผลการวิเคราะห์ตัวอย่างทำให้ได้ดีเอ็นเอที่สามารถนำมาเปรียบเทียบกับผลตรวจดีเอ็นเอในฐานข้อมูลได้ โดยมีหน่วยงานที่ใช้ประโยชน์จากข้อมูลดังกล่าวนี้มาเป็นระยะเวลายาวนาน และยังมีองค์กรหรือหน่วยงานที่เป็นสมาชิกประเทศต่าง ๆ อีกด้วย

เนื้อหา

1. การตรวจดีเอ็นเอในงานนิติวิทยาศาสตร์

การตรวจดีเอ็นเอในงานนิติวิทยาศาสตร์มีความสำคัญต่อกระบวนการยุติธรรม เนื่องจากผลการตรวจดีเอ็นเอเป็นหลักฐานทางวิทยาศาสตร์ที่ใช้พิสูจน์ตัวบุคคลเพื่อยืนยันตัวผู้กระทำความผิดซึ่งได้รับการยอมรับในระดับสากล โดยการตรวจดีเอ็นเอในห้องปฏิบัติการเป็นการวิเคราะห์ดีเอ็นเอลำดับเบสซ้ำต่อเนื่องแบบสั้นที่เรียกว่า STR (Short tandem repeated) หรือไมโครแซทเทลไลต์ (Microsatellite) ซึ่งมีความหลากหลายทางพันธุกรรมสูง วิธีการตรวจดีเอ็นเอมีหลายขั้นตอนและต้องใช้เครื่องมือในการวิเคราะห์หลายประเภทซึ่งมีค่าใช้จ่ายค่อนข้างสูง ดังนั้นจึงต้องพิจารณาการเก็บวัตถุพยานและสิ่งส่งตรวจเพื่อให้เกิดประโยชน์และความคุ้มค่าสูงสุด กระบวนการตรวจดีเอ็นเอในงานนิติวิทยาศาสตร์มีขั้นตอนการดำเนินการดังแสดงในภาพที่ 1 ซึ่งเป็นกระบวนการที่ใช้ความรู้ทางวิทยาศาสตร์หลายสาขา เช่น ชีววิทยา (Biology) ร่วมกับการใช้เทคโนโลยี (Technology) ในการตรวจดีเอ็นเอเพื่อให้ได้ผลตรวจดีเอ็นเอที่สามารถนำไปตรวจเปรียบเทียบเพื่อระบุตัวบุคคล ประกอบกับใช้ความรู้ด้านพันธุศาสตร์ (Genetics) และสถิติ (Statistics) ในการวิเคราะห์ความน่าจะเป็นด้านพันธุศาสตร์ (Probabilistic Genotyping) เพื่ออธิบายผลการตรวจดีเอ็นเอสำหรับใช้อ้างอิงในชั้นศาล ซึ่งวิธีที่ใช้มีหลายวิธีขึ้นอยู่กับห้องปฏิบัติการดีเอ็นเอว่าจะใช้วิธีใดในการวิเคราะห์ผลการตรวจเปรียบเทียบดีเอ็นเอ เช่น Random Match Probability (RMP)

เป็นวิธีที่แสดงความสัมพันธ์ของความน่าจะเป็นกับความถี่จากการสังเกตเหตุการณ์ที่เกิดขึ้นจากการทดลองหลายครั้ง หรือ Likelihood Ratio (LR) เป็นค่าสถิติที่ใช้ทดสอบเพื่อเปรียบเทียบค่าความน่าจะเป็นซึ่งเป็นอัตราส่วนระหว่างสมมติฐานระหว่างสมมุติฝ่ายโจทก์ (Prosecution hypothesis; H_p) และสมมุติฐานฝ่ายจำเลย (Defense Hypothesis; H_d) (Bulter, 2015)



ภาพที่ 1 การตรวจดีเอ็นเอในงานนิติวิทยาศาสตร์โดยใช้เครื่องหมายดีเอ็นเอ STR
ที่มา: (Bulter, 2015)

ผลการตรวจดีเอ็นเอจะถูกแสดงค่าเป็นตัวเลขของจำนวนซ้ำของเบสในแต่ละตำแหน่งของดีเอ็นเอ ในบทความนี้จะแสดงตัวอย่างผลตรวจดีเอ็นเอ 16 ตำแหน่ง โดยอ้างอิงจำนวนตำแหน่งที่กองพิสูจน์หลักฐานกลาง สำนักงานพิสูจน์หลักฐานตำรวจได้เลือกใช้ตั้งแต่เริ่มมีการตรวจพิสูจน์ดีเอ็นเอ ซึ่งปัจจุบันนี้ได้เพิ่มเติมจำนวนตำแหน่งดีเอ็นเออ้างอิงแล้ว แต่ตัวอย่างที่ปรากฏนี้นำมาจากข้อมูลดีเอ็นเอ PROVEDIT DATABASE ที่มีจำนวนตำแหน่งของดีเอ็นเอแบบเดียวกัน โดยบริษัทดังกล่าวได้เผยแพร่ข้อมูลชีวสารสนเทศของดีเอ็นเอโดยมีวัตถุประสงค์เพื่อใช้สำหรับการวิจัย (Alfonse et.al., 2018) สามารถดาวน์โหลดข้อมูลได้จากเว็บไซต์ ซึ่งมีข้อมูลดิบของผลตรวจดีเอ็นเอบุคคลเดี่ยวและดีเอ็นเอรูปแบบผสม (PROVEDIT Database LFTDI, 2019) ดังแสดงในภาพที่ 2 แสดงโครงสร้างข้อมูลลักษณะด้าเฟรมของผลตรวจดีเอ็นเอ 16 ตำแหน่ง ที่เป็นดีเอ็นเอรูปแบบเดี่ยว (Single source DNA profile) ดีเอ็นเอแต่ละตำแหน่งมีสองอัลลีลซึ่งประกอบด้วยตัวเลข 2 ชุด และตำแหน่งโครโมโซมเพศประกอบด้วยอักษรภาษาอังกฤษ X,X หรือ X,Y ในเพศหญิงหรือเพศชายตามลำดับ ต่อมาในปี ค.ศ. 2017 สำนักงานสอบสวนกลางของสหรัฐอเมริกาได้ประกาศให้ใช้ CODIS core เพิ่มขึ้นจำนวน 20 ตำแหน่ง (Karantzali et al., 2019) ในปัจจุบันตำแหน่งของดีเอ็นเอ STR ถูกเพิ่มเป็นจำนวน 24 ตำแหน่ง ซึ่งใช้กันอย่างแพร่หลาย (Gettings, Aponte, Vallone, & Butler, 2015)



	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	AM	D5S818	FGA
0	13,14	27,29	8,12	10,12	16,18	7,8	12,14	8,9	17,17	13,13	17,19	9,10	15,16	X,Y	12,12	22,24
1	14,14	30,30.2	8,9	12,12	15,18.2	7,9.3	11,13	9,11	22,24	11,13	15,17	8,8	14,17	X,Y	11,11	22,23
2	12,13	29,33.2	13,13	11,12	14,15	6,9	12,13	11,12	21,23	12,14	16,17	9,11	16,17	X,Y	12,12	20,24
3	11,12	28,33.2	8,10	10,12	15,17	6,8	9,12	10,10	22,23	11,13	17,17	6,10	18,20	X,Y	8,12	22,24
4	14,16	28,31.2	9,10	11,12	16,17	6,7	9,12	10,13	19,25	13,2,14.2	16,17	6,8	17,18	X,X	11,12	21,28
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
56	11,12	29,30	9,10	10,12	16,17	6,6	11,12	11,12	16,25	15,15	14,18	9,12	13,14	X,Y	11,13	22,23
57	8,14	28,29	8,12	10,12	15,16	8,9	10,11	8,13	23,24	13,15	14,16	8,11	13,15	X,X	10,12	21,23
58	14,14	30,30	9,11	10,11	16,17	8,9.3	11,12	10,13	19,23	14,15	17,17	8,11	14,17	X,X	12,13	21,23
59	13,14	28,29	10,11	12,14	14,15	9,9.3	12,12	8,11	16,18	14,15	17,17	8,8	15,15	X,Y	11,11	21,24
60	12,12	30,31	10,10	10,12	16,16	9,9.3	10,13	9,13	16,19	13,14	15,15	8,11	14,20	X,X	13,13	21,23

61 rows × 16 columns

ภาพที่ 2 ลักษณะข้อมูลผลตรวจดีเอ็นเอ STR จำนวน 16 ตำแหน่ง

หมายเหตุ : ผลลัพธ์ที่แสดงเกิดจากการเขียนไพทอนสคริปส์และแพนด้าดาต้าเฟรมเพื่อแสดงผลตัวอย่างข้อมูลดีเอ็นเอที่ถูกจัดเก็บ

2. หน่วยงานที่เป็นแหล่งข้อมูลด้านชีวสารสนเทศของดีเอ็นเอ

หน่วยงานหรือสถาบันที่เกี่ยวข้องกับชีวสารสนเทศของดีเอ็นเอเป็นแหล่งข้อมูลด้านชีวสารสนเทศที่สำคัญ ซึ่งมีส่วนเกี่ยวข้องกับด้านนิติวิทยาศาสตร์ เพื่อการวิจัย เพื่อทดลองและทดสอบงานของระบบ เพื่อเป็นแนวทางให้เห็นถึงความสำคัญที่แต่ละองค์กรได้ก่อตั้งขึ้นทั่วโลก แนวทางในการนำข้อมูลชีวสารสนเทศไปใช้ประโยชน์ดังแสดงในตารางที่ 1

ตารางที่ 1 องค์กรที่เป็นแหล่งข้อมูลชีวสารสนเทศเพื่อการวิจัยและอ้างอิง

หน่วยงาน	ลักษณะข้อมูล
1. สถาบันชีวสารสนเทศแห่งยุโรป (European Bioinformatics Institute) *	มี WebAPI ให้เลือกใช้เพื่อการวิจัยในส่วนที่เกี่ยวข้องกับดีเอ็นเอมนุษย์ เน้นข้อมูลลำดับนิวคลีโอไทด์ในจีโนมมนุษย์ (EMBL-EBI, 2021)
2. ศูนย์ข้อมูลเทคโนโลยีชีวภาพแห่งชาติ สหรัฐอเมริกา (NCBI) **	มีผลงานวิจัยและเครื่องมือด้านวิทยาศาสตร์และสุขภาพโดยให้การเข้าถึงข้อมูลด้านชีวการแพทย์และจีโนม มีฐานข้อมูลดีเอ็นเอชนิด Database of Short Genetic Variations (dbSNP) (Medicine, 2021)
3. สถาบันพันธุศาสตร์ ประเทศญี่ปุ่น (DNA Data Bank of Japan) *	แหล่งข้อมูลในระดับจีโนมเพื่อการวิจัยที่แบ่งปันกันเพื่อการวิจัยสำหรับการศึกษาที่สูงกว่าปริญญาตรี (Genetics, 2021)
4. สถาบันชีวสารสนเทศแห่งสวิสเซอร์แลนด์ (Swiss Institute of Bioinformatics (SIB)) *	ข้อมูลด้านดีเอ็นเอเพื่อการวิจัย ส่วนใหญ่มีไว้เพื่อการวิเคราะห์ดีเอ็นเอที่ผิดปกติในด้านการแพทย์และหาสาเหตุการเกิดโรค (Bioinformatics, 2021)



หน่วยงาน	ลักษณะข้อมูล
5. ศูนย์ชีวสารสนเทศแห่งชาติออสเตรเลีย (ARC Center in Informatics) *	ศูนย์วิจัยและพัฒนาชีวสารสนเทศการจัดการหาทรัพยากรด้านชีวสารสนเทศ ชีวสถิติ และแพลตฟอร์มการคำนวณเฉพาะทาง QFA มุ่งเน้นการวิจัยด้านชีวโมเลกุล (Ragan, 2006) (QFAB, 2021)
6. ศูนย์พันธุวิศวกรรมและเทคโนโลยีชีวภาพแห่งชาติ (BIOTEC) *	ศูนย์เก็บรวบรวมชีวสารสนเทศที่สนับสนุนงานด้านอาหารและการเกษตร การแพทย์ และสาธารณสุข พลังงานและสิ่งแวดล้อม ทรัพยากรชีวภาพ (BIOTEC, 2021)
7. ฐานข้อมูล STRBase สถาบันมาตรฐานและเทคโนโลยีแห่งชาติ (NIST) ***	รวบรวมผลงานตีพิมพ์มากมายเกี่ยวกับดีเอ็นเอในงานนิติวิทยาศาสตร์ มีเครื่องมือและทรัพยากรเพื่อวิเคราะห์หลายแบบโดยรวมถึงข้อมูลผลตรวจดีเอ็นเอ STR ที่มีให้ดาวน์โหลด (Butler & STRBase, 2008)
8. PROVEDIT DATABASE***	บริษัทเอกชนที่มีข้อมูลตัวอย่างซึ่งรวบรวมข้อมูลผลตรวจดีเอ็นเอ STR เพื่อการวิจัยให้กับนักวิจัยทั่วไป DATABASE (Alfonse et.al., 2018)

หมายเหตุ : * ข้อมูลชีวสารสนเทศ เน้นด้านจีโนม ชีวโมเลกุลของมนุษย์ และด้านการแพทย์

** ข้อมูลชีวสารสนเทศ เน้นด้านจีโนม ชีวโมเลกุลของมนุษย์ และด้านการแพทย์ รวมถึงฐานข้อมูลยีนแบบสั้น (dbSNP)

*** ชีวสารสนเทศของผลตรวจดีเอ็นเอ STR

แหล่งข้อมูลชีวสารสนเทศของดีเอ็นเอที่มีอยู่มากมายดังตัวอย่างองค์กรที่กล่าวมา เพื่อการนำไปวิจัยหรือเพื่ออ้างอิงการสร้างระบบของตนเอง ซึ่งในปัจจุบันได้ถูกผลักดันให้มีการเปลี่ยนไปใช้ฐานข้อมูล โดยสร้างการเข้าถึงและบูรณาการข้อมูลในฐานข้อมูลทางชีววิทยาที่เกี่ยวข้องกันหลายแห่ง เป็นความท้าทายสำคัญที่ผู้สร้างระบบฐานข้อมูลต้องสามารถเชื่อมโยงข้อมูลและความรู้จากแหล่งข้อมูลที่อยู่สถานที่ต่าง ๆ ทั่วโลก (Masys, 1989) เนื่องจากองค์กรด้านนิติวิทยาศาสตร์ส่วนใหญ่ใช้การตรวจดีเอ็นเอด้วยข้อมูลผลตรวจดีเอ็นเอ STR จึงมีหน่วยงานภายใต้สถาบันมาตรฐานและเทคโนโลยีแห่งชาติสหรัฐอเมริกา (NIST) ได้สร้าง Short Tandem Repeat DNA Internet DataBase หรือ STRBase ซึ่งเป็นฐานข้อมูลดีเอ็นเอในงานนิติวิทยาศาสตร์ที่ได้จัดตั้งในปี 1997 โดยสถาบันมาตรฐานดังกล่าวได้เก็บรวบรวมและบำรุงรักษาฐานข้อมูลดีเอ็นเอ STR ไว้บนอินเทอร์เน็ต ฐานข้อมูลนี้มีรายละเอียดเกี่ยวกับเครื่องหมายดีเอ็นเอ STR ที่ใช้กันทั่วไปและยังรวบรวมผลงานตีพิมพ์มากมายเกี่ยวกับดีเอ็นเอในงานนิติวิทยาศาสตร์ (Butler & STRBase, 2008) นอกจากนี้ยังมีข้อมูลเพื่อการวิจัยของบริษัทผู้ผลิตที่ได้เก็บรวบรวมไว้ให้ใช้เพื่อการวิจัยคือ PROVEDIT DATABASE (Alfonse et.al., 2018) ในบทความนี้จะขอใช้ข้อมูลตัวอย่างจากแหล่งที่กล่าวไปแล้วข้างต้นนี้



3. หน่วยงานกำหนดมาตรฐานข้อมูลดีเอ็นเอ

การใช้ชีวสารสนเทศกับผลตรวจดีเอ็นเอในงานนิติวิทยาศาสตร์นั้น ผลตรวจดีเอ็นเอที่ถูกเก็บต้องเป็นไปตามมาตรฐานการตรวจดีเอ็นเอทางนิติวิทยาศาสตร์ตามที่สถาบันมาตรฐานและเทคโนโลยีแห่งชาติสหรัฐอเมริการะบุ ซึ่งจะอยู่ในฐานข้อมูล STRBase (SRD-130) (Short Tandem Repeat DNA Internet DataBase) ที่ได้กำหนดโดยสถาบันมาตรฐานและเทคโนโลยีแห่งชาติสหรัฐอเมริกา (NIST Standard Reference Database) ต่อมาในปี 1995 เครือข่ายสถาบันนิติวิทยาศาสตร์แห่งยุโรป (European Network of Forensic Science Institutes; ENFSI) ได้ก่อตั้งขึ้น โดยมีวัตถุประสงค์เพื่อปรับปรุงการแลกเปลี่ยนข้อมูลร่วมกันในด้านนิติวิทยาศาสตร์ รวมถึงการปรับปรุงคุณภาพของข้อมูลที่ใช้ร่วมกันทางนิติวิทยาศาสตร์ในยุโรป เป็นจุดเด่นของเครือข่าย ENFSI จึงได้รับการยอมรับว่าเป็นองค์กรทางนิติวิทยาศาสตร์จากคณะกรรมการการยุโรป (ENFSI, 2021) นอกจากนี้สถาบันมาตรฐานและเทคโนโลยีแห่งชาติสหรัฐอเมริกายังให้ความสำคัญกับงานด้านการประกันคุณภาพ เนื่องจากเป็นสิ่งสำคัญที่ห้องปฏิบัติการทางนิติวิทยาศาสตร์ต้องดำเนินการ สถาบันมาตรฐานและเทคโนโลยีแห่งชาติสหรัฐอเมริกาจึงได้จัดตั้งคณะกรรมการที่ปรึกษาด้านมาตรฐานการประกันคุณภาพสำหรับห้องปฏิบัติการตรวจดีเอ็นเอทางนิติวิทยาศาสตร์ เพื่อเป็นแนวทางในการปฏิบัติด้านการประกันคุณภาพเกี่ยวกับมาตรฐานการประกันคุณภาพของห้องปฏิบัติการดีเอ็นเอ (Quality Assurance Standards for Forensic DNA Testing Laboratories) และมาตรฐานการประกันคุณภาพในการใช้ฐานข้อมูลสำหรับห้องปฏิบัติการดีเอ็นเอ (Quality Assurance Standards for DNA Databasing Laboratories) (Butler, DAB, SWGDAM, ENFSI & BSAG, 2021)

4. หน่วยงานที่ใช้ประโยชน์จากฐานข้อมูลชีวสารสนเทศของดีเอ็นเอ

การสร้างฐานข้อมูลชีวสารสนเทศของดีเอ็นเอ มีบทบาทสำคัญในการนำไปแก้ปัญหาอาชญากรรม เนื่องจากมีศักยภาพในการเชื่อมโยงกลุ่มอาชญากรรม และ/หรือ กำหนดตัวผู้ต้องสงสัยในที่เกิดเหตุ ที่สำคัญดีเอ็นเอสามารถช่วยพิสูจน์ความบริสุทธิ์ของบุคคล การระบุเหยื่อจากภัยพิบัติ และการค้นหาบุคคลสูญหาย โดยใช้วิธีการทางนิติวิทยาศาสตร์ ผลการวิเคราะห์ตัวอย่างทำให้ได้ผลตรวจดีเอ็นเอที่สามารถเปรียบเทียบกับผลตรวจดีเอ็นเอที่เก็บไว้ในฐานข้อมูลได้ โดยมีหน่วยงานที่ใช้ประโยชน์จากข้อมูลดังกล่าวนี้มาเป็นระยะเวลานาน ดังนี้

4.1 ฐานข้อมูลดีเอ็นเอขององค์กรตำรวจสากล (INTERPOL'S DNA database)

INTERPOL เป็นหน่วยงานตำรวจระดับนานาชาติ โดยดูแลฐานข้อมูลดีเอ็นเอระหว่างประเทศสมาชิก หน่วยงานตำรวจในแต่ละประเทศสมาชิกสามารถส่งผลตรวจดีเอ็นเอของผู้กระทำความผิด ผลตรวจดีเอ็นเอจากสถานที่เกิดเหตุ ผลตรวจดีเอ็นเอผู้สูญหาย และผลตรวจดีเอ็นเอศพที่ไม่ระบุชื่อไปค้นหยังฐานข้อมูลดีเอ็นเออัตโนมัติของ INTERPOL ฐานข้อมูลนี้เป็นที่รู้จักในชื่อ ดีเอ็นเอ Gateway เริ่มต้นในปี 2545 โดยมีดีเอ็นเอเพียงรายการเดียว แต่ภายในปี 2557 มีข้อมูลดีเอ็นเอ มากกว่า 150,000 รายการ จาก 73 ประเทศสมาชิก ประเทศที่เข้าร่วมได้ใช้ดีเอ็นเอ Gateway เป็นเครื่องมือในการสืบสวนคดีอาชญากรรมและตรวจพบการเชื่อมโยงที่อาจเกิดขึ้น ผลตรวจดีเอ็นเอที่ประเทศสมาชิกส่งมาสืบค้นเพื่อเชื่อมโยงข้อมูลระหว่างประเทศ มีจำนวนถึง 77 ประเทศ ในปี 2014 ประเทศสมาชิกสามารถเข้าถึงฐานข้อมูลผ่านระบบการสื่อสารของตำรวจสากล โดย INTERPOL ทำหน้าที่เป็นเพียงตัวกลางในการแบ่งปันและเปรียบเทียบข้อมูลเท่านั้น ผลตรวจดีเอ็นเอที่ถูกเก็บเป็นรูปแบบของชุดตัวเลขซึ่งแตกต่างกันในแต่ละบุคคล และไม่มีข้อมูลที่จะเชื่อมโยงว่าใครเป็นเจ้าของผลตรวจดีเอ็นเอนั้น ประเทศสมาชิกที่ใช้

ดีเอ็นเอ Gateway จะยังคงเป็นเจ้าของข้อมูลผลตรวจดีเอ็นเอของตนและไม่ขัดกับการทำลายล้าง (Disruption) ตามกฎหมายของประเทศนั้น ๆ (INTERPOL, 2019)

4.2 ฐานข้อมูลดีเอ็นเอสำนักงานสอบสวนกลางสหรัฐอเมริกา (FBI)

ระบบดัชนีดีเอ็นเอถูกสร้างและดูแลโดยสำนักงานสอบสวนกลางสหรัฐอเมริกา ประกอบด้วยข้อมูลสามระดับ ได้แก่ ระบบดัชนีดีเอ็นเอท้องถิ่น (LDIS) ระบบดัชนีดีเอ็นเอของรัฐ (SDIS) และระบบดัชนีดีเอ็นเอแห่งชาติ (NDIS) ผู้ตรวจสอบสามารถนำเข้าสู่ข้อมูลผลตรวจดีเอ็นเอและค้นหาการจับคู่กับผลตรวจกับดีเอ็นเอในคดีท้องถิ่น ผลตรวจดีเอ็นเอที่ถูกบันทึกในระดับท้องถิ่นจะถูกอัปโหลดหรือส่งไปยังระดับรัฐและระดับชาติ ซึ่งช่วยให้สามารถเปรียบเทียบข้อมูลดีเอ็นเอระหว่างกันได้ (FBI, 2008) ฐานข้อมูลของ CODIS มีฐานข้อมูลที่แตกต่างกันอยู่หลายฐานข้อมูล ขึ้นอยู่กับประเภทของข้อมูลที่ค้นหาฐานข้อมูลดีเอ็นเอในงานนิติวิทยาศาสตร์ ประกอบด้วย 1) ฐานข้อมูลดีเอ็นเอของบุคคล ได้แก่ ผู้กระทำความผิด ผู้ที่ถูกจับกุมในคดีอาญา บุคคลสูญหาย 2) ฐานข้อมูลของดีเอ็นเอที่เกี่ยวข้องกับสถานที่เกิดเหตุ 3) ฐานข้อมูลดีเอ็นเอศพนิรนาม ซึ่งซอฟต์แวร์ที่ใช้ในการตรวจเปรียบเทียบดีเอ็นเอของ CODIS สามารถตรวจเปรียบเทียบโดยตรง (Direct match) ว่าผลตรวจดีเอ็นเอเป็นของบุคคลเดียวกันหรือไม่ โดยความเสี่ยงในการจับคู่ระหว่างผลตรวจดีเอ็นเอในที่เกิดเหตุและผลตรวจดีเอ็นเอบุคคลบริสุทธิ์จะเพิ่มขึ้นเมื่อจำนวนผลตรวจดีเอ็นเอในฐานข้อมูลเพิ่มขึ้นและมีข้อมูลดีเอ็นเอในคดีต่าง ๆ มากขึ้น สามารถลดความเสี่ยงในการจับคู่ดีเอ็นเอในฐานข้อมูลดีเอ็นเอได้โดยการกำหนดเกณฑ์ขั้นต่ำ (Minimum Load Criterion; MLC) ในการจับคู่ดีเอ็นเอในฐานข้อมูลดีเอ็นเอ การตั้งค่า MLC ต่ำเกินไปจะเพิ่มความเสี่ยงของการจับคู่ ดีเอ็นเอผิดพลาด แต่การตั้งค่า MLC สูงเกินไปจะจำกัดจำนวนดีเอ็นเอที่สามารถเข้าสู่ฐานข้อมูลทำให้โอกาสในการจับคู่ดีเอ็นเอน้อยลง (Tvedebrink et.al, 2015) นอกจากนี้ระบบ CODIS สามารถค้นหา ดีเอ็นเอที่มีความสัมพันธ์แบบเครือญาติ (Familial searching) ในกรณีที่ไม่สามารถตรวจเปรียบเทียบโดยตรงได้เพื่อช่วยในการสืบสวนทางคดีให้แคบลง หรือการค้นหากรณีบุคคลสูญหาย หรือกรณีเกิดภัยพิบัติ เช่น การหาความสัมพันธ์แบบบิดา มารดาและบุตร หรือความสัมพันธ์แบบพี่น้อง เป็นต้น (Karantzali et al., 2019) ในปัจจุบันระบบฐานข้อมูลดีเอ็นเอแห่งชาติถูกนำมาใช้ในหลายประเทศทั่วโลก หลายประเทศอยู่ในช่วงการสร้างฐานข้อมูลดีเอ็นเอของตนเอง ซึ่งฐานข้อมูลดีเอ็นเอแต่ละประเทศอาจเครื่องหมายดีเอ็นเอ STR ที่เหมือนกัน หรือบางประเทศก็ใช้เครื่องหมายดีเอ็นเอ STR ที่แตกต่างจากฐานข้อมูลดีเอ็นเอของประเทศอื่นในบางตำแหน่ง

บทวิเคราะห์และวิจารณ์

1) การตรวจดีเอ็นเอในงานนิติวิทยาศาสตร์

การตรวจดีเอ็นเอในงานนิติวิทยาศาสตร์ เริ่มตั้งแต่วิธีการเก็บวัตถุพยานในสถานที่เกิดเหตุ การเก็บดีเอ็นเอจากตัวบุคคล การนำส่งวัตถุพยานมายังห้องปฏิบัติการตรวจดีเอ็นเอ ซึ่งเป็นไปตามหลักห่วงโซ่วัตถุพยาน (Chain of Custody) การเลือกใช้เครื่องมือหรือเทคโนโลยีในการตรวจดีเอ็นเอ เพื่อให้ได้ผลการตรวจออกมาเป็นที่ยอมรับเพื่อใช้อ้างอิงในชั้นศาล ในบทความนี้จะกล่าวเน้นแนวทางการนำข้อมูลผลการตรวจในทางปฏิบัติแล้วผลการตรวจนั้นเป็นข้อมูลชีวสารสนเทศชนิดเครื่องหมายดีเอ็นเอ STR ที่สามารถนำมาบูรณาการได้ทั้งด้านการวิจัย เพื่อนำไปสร้างระบบจัดเก็บไว้ในฐานข้อมูลชีวสารสนเทศของดีเอ็นเอ ตั้งแต่การนำข้อมูลนั้นมาวิเคราะห์ค่าทางสถิติ การเปรียบเทียบในรูปแบบต่าง ๆ และมีเครื่องมือที่สร้างขึ้นโดยองค์กรเพื่อใช้งานที่หลากหลายขึ้นอยู่กับว่าจะนำมาใช้งานในด้านใด แม้ว่าจะมีองค์กรหลักอย่าง



INTERPOL และ FBI ที่คอยผลักดันให้ใช้ข้อมูลนี้ร่วมกันแต่ก็มีบางองค์กรมองว่าผลการตรวจยืนยันในห้องปฏิบัติการประกอบด้วยข้อมูลสำคัญในการระบุตัวบุคคลที่ใช้ในการค้นหาและเปรียบเทียบดีเอ็นเอ นั้นเป็นข้อมูลส่วนบุคคลไม่สามารถเปิดเผยต่อที่สาธารณะได้ เช่น ประเทศไทยประกาศพระราชบัญญัติข้อมูลข่าวสารของราชการ พ.ศ. 2540 มาตรา 24 ได้กำหนด “ห้ามมิให้หน่วยงานของรัฐเปิดเผยข้อมูลข่าวสารส่วนบุคคลที่อยู่ในความดูแลของตนต่อหน่วยงานของรัฐแห่งอื่น หรือผู้อื่น โดยปราศจากความยินยอมเป็นหนังสือของเจ้าของข้อมูลที่ได้รับไว้ล่วงหน้าหรือในขณะนั้นมิได้” (Office of the Official Information Commission, 2008)

ความสำคัญของข้อมูลชีวสารสนเทศของดีเอ็นเอ นั้นเป็นสิ่งจำเป็น เพราะข้อมูลดังกล่าวสามารถนำมาสนับสนุนการทำงานของหน่วยงานที่เกี่ยวข้องด้านกระบวนการยุติธรรม การระบุตัวบุคคลได้ถูกต้อง ย่อมส่งผลต่อการคลี่คลายคดี อย่างไรก็ตามยังมีประเด็นอื่นที่เกี่ยวข้องอีกมากมาย การไม่ได้เข้าร่วมเป็นสมาชิก ที่สามารถใช้งานในระบบ CODIS ทั้งแง่ของนโยบายและกฎหมายของแต่ละองค์กร ในแต่ละประเทศทำให้เกิดช่องว่างที่แตกต่างในการใช้ข้อมูลชีวสารสนเทศ ในบทความนี้จะขอลำถึงข้อดีและข้อเสียของการเข้าร่วมและไม่เข้าร่วมเป็นสมาชิกขององค์กรที่สนับสนุนการใช้ระบบ CODIS ทั้ง INTERPOL และ FBI โดยแยกกรณีดังแสดงในตารางที่ 2 และไม่อาจกล่าวได้ว่าแบบใดดีกว่ากัน

ตารางที่ 2 การเปรียบเทียบข้อดีและข้อเสียกรณีการเข้าร่วมและไม่เข้าร่วมการใช้ระบบ CODIS

ประเด็น	ประเทศที่เป็นสมาชิก CODIS	ประเทศที่ไม่เป็นสมาชิก CODIS
1. การใช้งานระบบ	สมัครในฐานะองค์กรของรัฐที่เป็นสมาชิกและใช้ระบบได้	ไม่สามารถใช้งานระบบได้
2. การสร้างฐานข้อมูลชีวสารสนเทศของดีเอ็นเอ	ระบบมาตรฐานและสามารถรองรับการทำงานสำหรับประเทศสมาชิกทั่วโลก	องค์กรต้องสร้างฐานข้อมูลชีวสารสนเทศของดีเอ็นเอขึ้นเอง และใช้เฉพาะหน่วยงานนั้น
3. ปริมาณข้อมูล	มีปริมาณข้อมูลมากและหลากหลายตามกาลเวลา ตั้งแต่อดีตถึงปัจจุบันมีการพัฒนาอย่างต่อเนื่อง เช่น ตำแหน่งดีเอ็นเอที่อ้างอิงต่างกัน การทำดัชนีแยกกระหว่างคดีอาชญากรรมหรือบุคคลสูญหาย เป็นต้น	ปริมาณข้อมูลมีจำกัดเฉพาะในองค์กรขึ้นอยู่กับจำนวนคดีที่เกิดขึ้น
4. การแบ่งปันข้อมูล	มีการแบ่งปันข้อมูลตามรูปแบบมาตรฐานระหว่างประเทศ ทำให้การคลี่คลายคดีทำได้ในวงกว้างขึ้น เช่น คดีอาชญากรรมระหว่างประเทศ เป็นต้น	ไม่สามารถแบ่งปันข้อมูลได้ เนื่องจากองค์กรมีนโยบายตั้งแต่ต้นว่าข้อมูลชีวสารสนเทศของดีเอ็นเอ นั้นเป็นความลับส่วนบุคคลไม่สามารถเปิดเผยได้
5. ความถูกต้องและรวดเร็ว	ระบบนี้ถึงจะมีจำนวนข้อมูลมหาศาล การเข้าถึงข้อมูลสามารถขุดเซยเทคโนโลยีที่ทันสมัยมากขึ้น และด้วย	สามารถเข้าถึงได้ ถูกต้องรวดเร็ว เนื่องจากมีปริมาณข้อมูลน้อย ทั้งนี้ขึ้นอยู่กับวิธีการออกแบบและสร้าง



ประเด็น	ประเทศที่เป็นสมาชิก CODIS	ประเทศที่ไม่เป็นสมาชิก CODIS
	วิธีการทำดัชนี เพื่อการเข้าถึงข้อมูลที่ถูกต้อง รวดเร็ว และรองรับการทำงานระหว่างประเทศสมาชิก	ระบบว่าข้อมูลที่ใส่ลงไปจะมีการจัดการอย่างถูกต้องเพียงใด
6. มาตรฐาน	องค์กรหลักดังกล่าวเป็นทั้งองค์กรที่สร้างระบบและในฐานะที่เป็นองค์กรมาตรฐานรองรับการทำงานมาเป็นเวลายาวนานและเป็นที่ยอมรับ	ในรายละเอียดแม้จะอ้างอิงมาตรฐานของสถาบันมาตรฐานและเทคโนโลยีแห่งชาติสหรัฐอเมริกา แต่การออกแบบฐานข้อมูลอาจมีความหลากหลายเชิงเทคนิค เช่น บางองค์กรอาจไม่ออกแบบข้อมูลที่เป็นบรรทัดฐาน (Normalization) หรือออกแบบตามความต้องการใช้ข้อมูลของแต่ละองค์กรเท่านั้น
7. นโยบายความเป็นส่วนตัว	แม้องค์กรดังกล่าวจะบอกว่าเป็นเพียงการส่งข้อมูลเพื่อการเปรียบเทียบ ไม่เกี่ยวกับข้อมูลส่วนบุคคลอื่น ๆ แต่ก็ยังคงเกิดความไม่มั่นใจของผู้ใช้	สนับสนุนนโยบายส่วนบุคคลได้อย่างดีเพราะว่าไม่มีการแบ่งปันข้อมูลแต่อย่างใด

ระดับนโยบายขององค์กรมีผลต่อองค์กรอย่างมาก ดังนั้นกำหนดนโยบายการบริหารและพัฒนาขององค์กรด้านชีวสารสนเทศต้องชัดเจนจึงจะมีทิศทางในการปฏิบัติ องค์กรในหลายประเทศยังไม่มีผลกระทบต่อการเข้าร่วมใช้ระบบ CODIS แต่จะสร้างระบบขึ้นมาด้วยตนเอง ซึ่งอาจเป็นเหตุผลในระดับนโยบายขององค์กรของแต่ละประเทศที่ยังไม่สามารถเข้าร่วมระบบ CODIS ได้

2) หน่วยงานที่เป็นแหล่งข้อมูลด้านชีวสารสนเทศของดีเอ็นเอ

เนื่องจากความเป็นข้อมูลส่วนบุคคล ในระยะการสร้างระบบการจัดเก็บฐานข้อมูลชีวสารสนเทศของดีเอ็นเอนั้นจะต้องมีข้อมูลในการทดลองและทดสอบระบบ การจะนำข้อมูลจริงในห้องปฏิบัติการมาใช้ในการออกแบบและทดสอบนั้น ผิดต่อหลักการในเรื่องการเปิดเผยข้อมูลส่วนบุคคล จึงต้องหาแหล่งข้อมูลชีวสารสนเทศของดีเอ็นเอที่เปิดเผยมาใช้ ซึ่งองค์กรดังกล่าวนี้มีข้อมูลเพื่อสนับสนุนการใช้งาน และเพื่อการวิจัย พิจารณาจากตารางที่ 1 พบว่าส่วนใหญ่แล้วจะมุ่งเน้นไปที่ชีวโมเลกุลของจีโนมมนุษย์ที่สำคัญต่อวงการแพทย์ ส่วนข้อมูลดีเอ็นเอ STR ที่สนับสนุนงานด้านนิติวิทยาศาสตร์นั้นมีเพียง 2 องค์กรที่เน้นข้อมูลจากผลการตรวจดีเอ็นเอในงานนิติวิทยาศาสตร์ซึ่งมุ่งไปที่การตรวจเพื่อยืนยันตัวบุคคล ดังนั้นองค์กรที่ใช้เทคโนโลยีด้านนี้ซึ่งมีอยู่ทั่วโลกแต่จะใช้เฉพาะกลุ่มนักนิติวิทยาศาสตร์ ซึ่งต่างจากการตรวจในด้านทางการแพทย์ที่วิเคราะห์หาการเกิดโรคหรือความผิดปกติของดีเอ็นเอ ในบางประเทศมองว่าการกระทำผิดต่าง ๆ อาจเกิดจากภาวะการเกิดโรคที่มีความผิดปกติของดีเอ็นเอ ซึ่งในทีนี้จะไม่กล่าวไปถึงหลักอาชญาวิทยาและพันธุวิทยา

กรณีการไม่เข้าร่วมใช้ระบบ CODIS และต้องสร้างระบบขึ้นมาเอง ผู้สร้างควรรู้แหล่งข้อมูลที่เป็นประโยชน์ที่นักวิจัยทั่วโลกอ้างอิง และมีความเป็นมาตรฐานดังได้แสดงไว้ในตารางที่ 1 ในส่วนที่เกี่ยวข้องกับนิติวิทยาศาสตร์ ต้องพิจารณาแหล่งข้อมูลที่สนับสนุนระบบการตรวจพิสูจน์ดีเอ็นเอรูปแบบเดียวกัน เช่น



การเลือกผลการตรวจดีเอ็นเอด้วยเครื่องหมาย STR ควรเลือกข้อมูลดีเอ็นเอชนิดเครื่องหมาย STR จาก STRBase หรือ PROVEDIT DATABASE ถ้าเป็นการตรวจพิสูจน์บุคคลโดยใช้สลิป Single Nucleotide Polymorphisms (SNPs) ควรเลือก NCBI เป็นต้น แต่ทั้งหมดนี้จะต้องคำนึงถึงการออกแบบและสร้างระบบให้สอดคล้องกับระบบหลักสำหรับการรองรับงานในอนาคตและที่สำคัญข้อมูลดังกล่าวเมื่อนำไปใช้เพื่อการวิจัยหรือทดลองใด ๆ จะต้องไม่ผิดหลักของจรรยาบรรณของการวิจัย

3) หน่วยงานกำหนดมาตรฐานข้อมูลดีเอ็นเอและหน่วยงานที่ใช้ประโยชน์จากฐานข้อมูลดีเอ็นเอ

องค์กรที่ไม่ได้เข้าร่วมสมาชิกระบบดัชนีดีเอ็นเอ จะต้องจัดสร้างระบบการจัดเก็บฐานข้อมูลชีวสารสนเทศของดีเอ็นเอขึ้นมาเอง และหลีกเลี่ยงไม่ได้ว่าจะต้องมีมาตรฐานเบื้องต้นกำกับด้วยองค์การอ้างอิงซึ่งเป็นที่ยอมรับกันทั่วโลก คือ สถาบันมาตรฐานและเทคโนโลยีแห่งชาติสหรัฐอเมริกา ดังนั้นการสร้างระบบจัดเก็บชีวสารสนเทศจะต้องอ้างอิงมาตรฐาน นอกจากนี้ยังมีเครือข่ายสถาบันนิติวิทยาศาสตร์แห่งยุโรป (European Network of Forensic Science Institutes ; ENFSI) ได้ก่อตั้งขึ้นโดยมีวัตถุประสงค์เพื่อปรับปรุงการแลกเปลี่ยนข้อมูลร่วมกันในด้านนิติวิทยาศาสตร์ รวมถึงการปรับปรุงคุณภาพของข้อมูลที่ใช้ร่วมกันทางนิติวิทยาศาสตร์ในยุโรป และกลายมาเป็นองค์กรที่ได้รับการยอมรับในด้านการปฏิบัติและการประกันคุณภาพห้องปฏิบัติการด้านนิติวิทยาศาสตร์ ทั้งนี้ต้องกำหนดข้อมูลสำคัญตามมาตรฐานสถาบันมาตรฐานและเทคโนโลยีแห่งชาติสหรัฐอเมริกาใน STRBase (SRD-130)

เนื่องจากองค์กรที่กำหนดมาตรฐานนั้นได้กำหนดไว้เป็นแนวทางเป็นโครงร่างมาตรฐาน แต่การสร้างระบบอาจมีกำหนดเสริมมากขึ้นกว่าเดิม การออกแบบจะเป็นอิสระตามประสบการณ์ของนักออกแบบฐานข้อมูล จึงเกิดความหลากหลายของรูปแบบการจัดเก็บข้อมูล การทดสอบความถูกต้องกรณีที่มีจำนวนข้อมูลมากขึ้น ความครอบคลุมการสืบค้นข้อมูลในฐานข้อมูลดีเอ็นเอ กรณีการเปรียบเทียบข้อมูลชีวสารสนเทศของดีเอ็นเอที่ใช้เครื่องหมายดีเอ็นเอ STR ต่างกัน อาจมีข้อดีและข้อด้อยต่างกัน กรณีองค์กรหรือประเทศที่ไม่เป็นสมาชิกของระบบ CODIS จะไม่สามารถค้นหาหรือแบ่งปันข้อมูลสำหรับการค้นหาเพื่อเปรียบเทียบผลตรวจดีเอ็นเอระหว่างประเทศได้ แต่ถึงอย่างไรรูปแบบข้อมูลผลตรวจดีเอ็นเอที่จัดเก็บไว้ในฐานข้อมูลดีเอ็นเอต้องเป็นไปตามมาตรฐานที่สถาบันมาตรฐานและเทคโนโลยีแห่งชาติสหรัฐอเมริกาคำหนด ดังนั้นถ้าต้องเปลี่ยนไปใช้ระบบ CODIS อาจต้องดัดแปลงข้อมูลผลตรวจดีเอ็นเอที่มีลำดับการจัดเรียงของเครื่องหมายดีเอ็นเอ STR ตามประเภทของชุดน้ำยาที่ใช้ในการตรวจดีเอ็นเอ ด้วยวิธีการที่ยุ่งยากเพื่อตรวจเปรียบเทียบผลตรวจดีเอ็นเอ นอกจากนี้การดูแลบำรุงรักษาระบบฐานข้อมูล ต้องมีระบบสำรองข้อมูลที่ครอบคลุมและมีผู้เชี่ยวชาญดูแลอย่างต่อเนื่อง ทั้งหมดนี้คือค่าใช้จ่ายที่ต้องบริหารจัดการเนื่องจากวงจรชีวิตของซอฟต์แวร์ที่สร้างขึ้นนั้นย่อมต้องมีการแก้ไขเมื่อมีความผิดปกติในการใช้งาน หรือพบเจอสภาพการใช้งานที่จำกัด จึงจำเป็นต้องมีการแก้ไขอย่างต่อเนื่อง

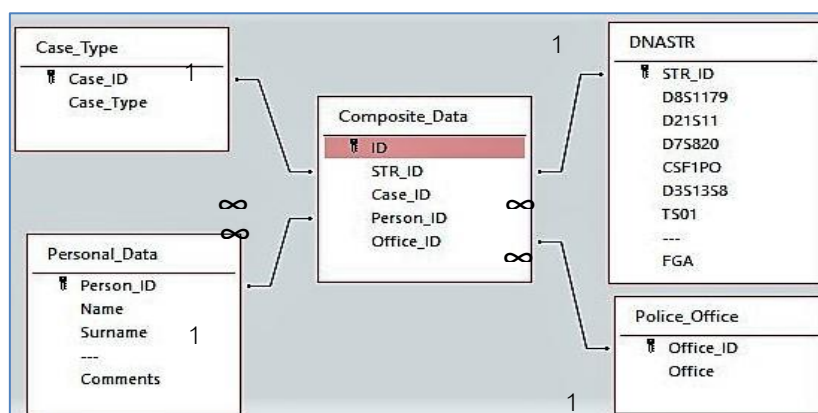
เนื่องจากการออกแบบระบบมีส่วนสำคัญอย่างยิ่งต่อการทำงาน นอกจากจะอ้างอิงมาตรฐานของสถาบันมาตรฐานและเทคโนโลยีแห่งชาติสหรัฐอเมริกาแล้ว จะต้องคำนึงถึงรายละเอียดตามความต้องการของผู้ใช้ระบบ ลักษณะข้อมูลที่ได้จากเครื่องตรวจวิเคราะห์ดีเอ็นเอนั้น จะมีลักษณะเป็นข้อมูลรายการบันทึกในรายละเอียด (Transection Data) เป็นแฟ้มส่วนขยาย .fsa (Fragment Size Analysis) ซึ่ง FSA เป็นไฟล์ข้อมูลในการวิเคราะห์ชิ้นส่วน ที่สร้างโดยเครื่องวิเคราะห์ดีเอ็นเออัตโนมัติซึ่งเป็นอิเล็กทรอนิกส์แบบแคพิลลารี (Capillary electrophoresis; CE) ที่สามารถหาลำดับนิวคลีโอไทด์ (Sequencing) และวิเคราะห์ชิ้นส่วนดีเอ็นเอ (Fragment analysis) จากนั้นนำไฟล์ไปวิเคราะห์ด้วยโปรแกรมสำหรับวิเคราะห์ข้อมูล ตัวอย่างโปรแกรมที่ใช้ในการวิเคราะห์ข้อมูล คือ Gene Mapper ID-X เพื่อสร้างอิเล็กทรอนิกส์

(Electropherogram; EPG) จะมีการบันทึกอิเล็กโตรฟีโรแกรมที่เวลาเริ่มต้นและสิ้นสุดของไฟล์ อิเล็กโตรฟีโรแกรมที่สร้างขึ้นประกอบด้วยแกน X แสดงเวลาที่ตีเอ็นเอแต่ละขนาดวิ่งออกมา และแกน Y แสดงปริมาณความเข้มข้นของสารเรืองแสงที่เครื่องวัดได้ โดยจะถูกบันทึกในรูปแบบไฟล์ ABIF ที่เป็นลิขสิทธิ์ของบริษัท ซึ่งแต่ละองค์ประกอบของข้อมูลเชื่อมโยงกับเครื่องหมาย (tag) ที่กำหนด โดยแต่ละห้องปฏิบัติการควรกำหนดเกณฑ์ที่ใช้ในการแปลผลการตรวจตีเอ็นเอให้เป็นไปตามเกณฑ์ Scientific working group on DNA analysis method (SWG DAMS) (Rerkamnuaychoke, 2019)

Sample File	Marker	Dye	Allele 1	Size 1	Height 1	Allele 2	Size 2	Height 2	Allele 3	Size 3	Height 3	Allele 4	Size 4	Height 4	Allele 5	Size 5	Height 5	Allele 6	Size 6	Height 6	Allele 7	Size 7	Height 7	
D05_RD14-0003-01d3a-0.0078F8C-Q0.7_04.25sec.hid	AMEL	B	OL	84.04	4																			
D05_RD14-0003-01d3a-0.0078F8C-Q0.7_04.25sec.hid	D3S1358	B	OL	116.06	7	17	130.96	29	18	135.23	365	OL	145.35	8										
D05_RD14-0003-01d3a-0.0078F8C-Q0.7_04.25sec.hid	D1S1656	B	OL	160.51	6	OL	178.98	5	15	181.8	100	17	190.04	8	18	194.19	7							
D05_RD14-0003-01d3a-0.0078F8C-Q0.7_04.25sec.hid	D2S441	B		10	219.67	18	11	223.75	273	12	228.13	9	OL	228.94	8	OL	237.03	7	OL	241.09	5			
D05_RD14-0003-01d3a-0.0078F8C-Q0.7_04.25sec.hid	D10S1248	B	OL	264.23	10	11	267.63	6	12	271.5	7	13	275.36	51	14	279.44	525	17	291.75	6				
D05_RD14-0003-01d3a-0.0078F8C-Q0.7_04.25sec.hid	D13S317	B	OL	308.24	4	6	309.19	9	7	313.16	18	8	317.05	166	11	329.26	12	12	333.44	152	OL	341.12	5	
D05_RD14-0003-01d3a-0.0078F8C-Q0.7_04.25sec.hid	Penta E	B	OL	361.64	3	OL	376.89	8	OL	381.46	7	OL	386.15	10	OL	391.66	11	11	396.98	250	OL	405.82	6	
D05_RD14-0003-01d3a-0.0078F8C-Q0.7_04.25sec.hid	D16S539	G	OL	86.14	11	OL	88.49	10	10.3	104.32	8	12	109.84	195	13	114.27	135							
D05_RD14-0003-01d3a-0.0078F8C-Q0.7_04.25sec.hid	D18S51	G		11.2	149.86	9	OL	151.3	18	13	156.5	158	19.2	183.08	10	22	193.5	13	OL	202.35	7	25	205.67	6
D05_RD14-0003-01d3a-0.0078F8C-Q0.7_04.25sec.hid	D2S1338	G		13	234.32	11	14	238.47	11	OL	243.99	11	OL	255.69	11	19	258.72	21	20	262.76	158	OL	285.81	12

ภาพที่ 3 แสดงข้อมูลรายละเอียด (Transection) โดยจะต้องแยกข้อมูลเพื่อนำไปออกแบบตามจุดประสงค์และความต้องการของผู้ใช้งานระบบ

จากภาพที่ 3 ข้อมูลผลการตรวจตีเอ็นเอมีจำนวนมากที่ถูกบันทึกไว้ แต่ไม่สามารถแสดงได้หมดจึงจำเป็นต้องแยกออกมาตามลักษณะเครื่องหมายตีเอ็นเอที่ห้องปฏิบัติการอ้างอิง ในบทความนี้จะเลือกแยกผลตามเครื่องหมายตีเอ็นเอ 16 ตำแหน่ง จะต้องแยกข้อมูลออกมาเพื่อทำการออกแบบโครงสร้างข้อมูลเป็นตารางในฐานข้อมูล ภายใต้เงื่อนไขและทำข้อมูลให้อยู่ในขั้นบรรทัดฐาน (Normalization) การสร้างดัชนีและคีย์หลัก การออกแบบฐานข้อมูลมีความหลากหลายในระดับแนวคิดในการสร้าง การจัดเก็บและนำไปใช้งาน แต่ต้องกำหนดข้อมูลสำคัญตามมาตรฐานสถาบันมาตรฐานและเทคโนโลยีแห่งชาติสหรัฐอเมริกาที่ระบุใน STRBase (SRD-130) และการออกแบบภายใต้เงื่อนไขและทำข้อมูลให้อยู่ในขั้นบรรทัดฐาน (Normalization)



ภาพที่ 4 มุมมองแนวคิดการเก็บข้อมูลผลการตรวจตีเอ็นเอในฐานข้อมูลชีวสารสนเทศของตีเอ็นเอ



ประเด็นการสร้างดัชนีและคีย์หลักนั้น ผู้เขียนบทความมีเห็นว่าต้องแยกส่วนเฉพาะดีเอ็นเอกับข้อมูลส่วนบุคคลอื่น ๆ แต่จะอ้างอิงถึงได้ด้วยคีย์ร่วมในตารางรวม (Composite Table) จะสามารถรองรับในอนาคตถ้าพบว่าสามารถนำข้อมูลไปใช้ประโยชน์ภายใต้กฎหมายระบุ ก็จะสามารถส่งออกข้อมูลเพื่อแบ่งปันได้ทันที ดังภาพที่ 4 ตัวอย่างที่นำเสนอนี้เป็นเฉพาะเพียงมุมมองแนวคิด (Conceptual View) จากการวิเคราะห์และออกแบบจากรายละเอียดข้อมูล (Transaction) เพื่อให้เห็นตัวอย่าง ซึ่งในความเป็นจริงอาจมีรายละเอียดมากกว่านี้ตามการออกแบบเพื่อการใช้งานของแต่ละองค์กร

บทสรุปและข้อเสนอแนะ

1) บทสรุป

ผู้เขียนบทความแสดงให้เห็นทราบความเป็นมาของหน่วยงานที่เกี่ยวข้องและการอ้างอิงมาตรฐานตลอดจนการนำข้อมูลที่ได้ไปใช้ประโยชน์ในเชิงการวิจัยก่อนที่จะนำมาออกแบบและสร้างระบบเพื่อนำไปใช้งาน ปัจจุบันเทคโนโลยีได้มีบทบาทในการนำมาใช้ประโยชน์และเพิ่มประสิทธิภาพการทำงาน ผู้เขียนบทความแสดงให้เห็นถึงการใช้ประโยชน์จากผลตรวจดีเอ็นเอกับเทคโนโลยีฐานข้อมูลที่สามารถจัดเก็บชีวสารสนเทศโดยผลการตรวจดีเอ็นเอด้วยเครื่องหมาย STR ในงานนิติวิทยาศาสตร์ แนวคิดการออกแบบและวิเคราะห์การสร้างฐานข้อมูลชีวสารสนเทศของดีเอ็นเอ จากผลลัพธ์ของเครื่องตรวจวิเคราะห์โดยอ้างอิงมาตรฐานและออกแบบไว้ในอนาคต สำหรับการส่งออกและแบ่งปันข้อมูล ตลอดจนการแลกเปลี่ยนข้อมูลผลตรวจดีเอ็นเอด้วยเครื่องหมาย STR ตามข้อกำหนดมาตรฐานข้อมูลดีเอ็นเอขององค์กรที่กำหนดมาตรฐาน

2) ข้อเสนอแนะ

ข้อเสนอแนะระดับนโยบายในเรื่องการจัดการฐานข้อมูลและการเข้าร่วมเป็นสมาชิกระบบ CODISควรมีความชัดเจนและมีตัวบทกฎหมายรับรอง จึงจะทำให้ทิศทางการทำงานด้านข้อมูลชีวสารสนเทศของ ดีเอ็นเอชัดเจนขึ้น รวมถึงการนำข้อมูลชีวสารสนเทศมาใช้ประโยชน์ควรเลือกแหล่งข้อมูลชีวสารสนเทศของดีเอ็นเอให้ตรงตามความต้องการใช้งาน เช่น การใช้งานด้านนิติวิทยาศาสตร์ ควรเลือกข้อมูลชีวสารสนเทศของดีเอ็นเอเครื่องหมาย STR เป็นต้น ซึ่งหากองค์กรต้องจัดทำระบบฐานข้อมูลดีเอ็นเอเพื่อใช้ในองค์กร ระบบฐานข้อมูลที่จัดทำขึ้นควรมีคุณลักษณะเพื่อรองรับการใช้งานของระบบ CODIS ดังนี้

- (1) โปรแกรมการเปรียบเทียบข้อมูลต้องรองรับการจัดเก็บผลตรวจดีเอ็นเอได้จำนวนมาก
- (2) ระบบจัดเก็บและการจัดการข้อมูลดีเอ็นเอต้องสามารถนำเข้าข้อมูลผลตรวจดีเอ็นเอจากเครื่องวิเคราะห์ดีเอ็นเอด้วยโปรแกรมการวิเคราะห์รูปแบบดีเอ็นเอห้องปฏิบัติการใช้ โดยสามารถวิเคราะห์ข้อมูลดิบ (Raw data) ในรูปแบบ .fsa และ .hid จากเครื่องวิเคราะห์ดีเอ็นเอได้
- (3) ระบบสามารถรองรับการใช้งานเมื่อเพิ่มตำแหน่งของเครื่องหมายดีเอ็นเอ STR
- (4) ระบบสามารถจับคู่ผลตรวจดีเอ็นเอแบบเปรียบเทียบโดยตรง (Direct match) วิเคราะห์ดีเอ็นเอในรูปแบบผสม (Mixture DNA profile) และค้นหาดีเอ็นเอที่มีความสัมพันธ์แบบเครือญาติ (Familial searching) ได้
- (5) ระบบสามารถวิเคราะห์ความน่าจะเป็นด้านพันธุศาสตร์ (Probabilistic Genotyping) เช่น Random Match Probability (RMP) ค่า Likelihood Ratio (LR) สามารถคำนวณค่าความสัมพันธ์ทางสายโลหิต (Kinship analysis) และความสัมพันธ์แบบพ่อ-แม่-ลูก (Paternity Index) ได้

(6) ระบบสามารถคำนวณข้อมูลความถี่อัลลีล (Allele frequency) และฐานข้อมูลประชากร (Population database) ได้

(7) มีระบบการบริหารจัดการเกี่ยวกับความปลอดภัยของข้อมูล

(8) การดำเนินการระบบฐานข้อมูลในขั้นตอนเริ่มต้นไม่ควรนำข้อมูลจริงมาทดลองหรือทดสอบเพราะอาจทำให้เกิดผลกระทบได้ เนื่องจากค่าตำแหน่งเครื่องหมายดีเอ็นเอเปลี่ยนไปย่อมมีผลกับเอกลักษณ์บุคคล เมื่อทดสอบระบบจนใช้งานได้ดีแล้วจึงจะนำข้อมูลจริงมาใช้

เอกสารอ้างอิง

- Alfonse, L. E., Garrett, A. D., Lun, D. S., Duffy, K. R. and Grgicak, C. M. (2018). A large-scale dataset of single and mixed-source short tandem repeat profiles to inform human identification strategies: PROVEDIt. **Forensic Science International: Genetics**, 32, 62-70.
- Bioinformatics, S. I. (2021). **Swiss Institute of Bioinformatics**. Retrieved July 5, 2021. from <https://www.sib.swiss/about-sib/organization>
- BIOTEC. (2021). **BIOTEC NSTDR**. Retrieved June 28, 2021. from <https://www.biotec.or.th/th/index.php>
- Bulter, J. M. (2015). **Advanced Topics in Forensic DNA Typing: Interpretation**. USA: Elsevier Inc.
- Butler, J. M. and STRBase (2008). **Forensic Science International: Genetics Supplement Series**, 1(1), 97-99. Retrieved July 3, 2021. from <https://sciencedirect.com/science/article/pii/S187517680800070x>.
- Butler, J. M., DAB, SWGDAM, ENFSI, and BSAG. (2021). **Forensic DNA Advisory Groups**. Retrieved July 3, 2021. from <https://nist.gov/publications/forensic-dna-advisory-groups-dab-swgdam-enfsi-and-bsag>.
- EMBL-EBI. (2021). **EMBL European Bioinformatics Institute**. Retrieved July 3, 2021. from Embl-Ebi: <http://www.ebi.ac.uk>
- ENFSI. (2021). **Welcome to ENFSI**. Retrieved July 5, 2021. from ENFSI: <http://www.enfsi.eu/>
- FBI. (2008). **FBI Handbook of Crime Scene Forensics**. New York: Skyhorse Publishing.
- Genetics, N. I. (2021). **National Institute of Genetics**. Retrieved July 7, 2021. from <https://www.nig.ac.jp/nig/>.
- Gettings, K. B., Aponte, R. A., Vallone, P. M. and Butler, J. M. (2015). STR allele sequence variation: Current knowledge and future issues. **Forensic Science International-genetics**, 18, 118-130. Retrieved July 3, 2021. from <https://nist.gov/publications/str-allele-sequence-variation-current-knowledge-and-future-issues>.
- INTERPOL. (2019). **Engaging Co-creation for Future Security Threats Final Report**. Lyon: Interpol Innovation Centre|@INTERPOL_GCI.
- Karantzali, E., Rosmaraki, P., Kotsakis, A., Pajolec, M.R. and Fitsialos, G. (2019). The effect of FBI CODIS Core STR Loci expansion on familial DNA database searching. **Forensic Science International: Genetics**, 43, 102129.



- Masys, D. R. (1989). New directions in bioinformatics. **Journal of Research of the National Institute of Standards and Technology**, 94(1), 59-63. Retrieved July 3, 2021. from <http://europepmc.org/abstract/med/28053399>.
- Medicine, U. S. (2021). **National Center for Biotechnology Information**. Retrieved July 7, 2021. from NCBI: <https://www.ncbi.nlm.nih.gov/>.
- Office of the Official Information Commission. (2008). **State agency's authorities and duties in protecting personal information**. Retrieved July, 5 2021. from http://www.oic.go.th/web2017/disclosure_pi_without_consent_data_subject.htm (In Thai).
- PROVEDIT Database LFTDI. (2019). **LFTID Laboratory for Forensic Technology Development & Integration**. Retrieved July 6, 2021. from <https://lftdi.camden.rutgers.edu/provedit/files/>.
- QFAB. (2021). **QFAB Bioinformatics**. Retrieved July 6, 2021. from <https://qfab.org/>
- Ragan, M. A. (2006). ARC CENTER IN BIOINFORMATICS. **Asia-pacific Biotech News**, 10(24), 1391-1394. Retrieved July 3, 2021. from <https://espace.library.uq.edu.au/view/uq:167210>.
- Rerkamnuaychoke, B. (2019). **Recent Advances in Forensic Genetics**. Bangkok: Beyond Enterprise. (In Thai).
- Tvedebrink, T., Bright J. A., Buckleton, J. S. and Curran, J. M. (2015). The effect of wild card designations and rare alleles in forensic DNA database searches. **Forensic Science International: Genetics**, 16, 98-104.

ประวัติผู้เขียน

คำนำหน้า ชื่อ-สกุล	พันตำรวจโทหญิง พรวิภา เวฬุภาณูญา *
ตำแหน่ง/สถานะ	อาจารย์ (สัญญาบัตร 3)
ที่อยู่หน่วยงาน/สังกัด	กลุ่มงานคณาจารย์ คณะนิติวิทยาศาสตร์ โรงเรียนนายร้อยตำรวจ 90 หมู่ 7 ตำบลสามพราน อำเภอสามพราน จังหวัดนครปฐม 73110
ไปรษณีย์อิเล็กทรอนิกส์	pornvipa@rpca.ac.th
คำนำหน้า ชื่อ-สกุล	พันตำรวจตรีหญิง สุพรรณษา มะลิขาว
ตำแหน่ง/สถานะ	อาจารย์ (สัญญาบัตร 2)
ที่อยู่หน่วยงาน/สังกัด	กลุ่มงานคณาจารย์ คณะนิติวิทยาศาสตร์ โรงเรียนนายร้อยตำรวจ 90 หมู่ 7 ตำบลสามพราน อำเภอสามพราน จังหวัดนครปฐม 73110
ไปรษณีย์อิเล็กทรอนิกส์	supunsarpca@hotmail.com

* ผู้ประพันธ์บรรณกิจ (Corresponding Author)