

PEMANFAATAN TEKNOLOGI SEKUENSING GENOM UNTUK MEMPERCEPAT PROGRAM PEMULIAAN TANAMAN

Utilization of Genome Sequencing Technology to Accelerate Plant Breeding Program

I Made Tasma

Balai Besar Penelitian dan Pengembangan Bioteknologi dan Sumberdaya Genetik Pertanian
Jalan Tentara Pelajar No. 3A Bogor 16111, Indonesia
Telp. (0251) 8339793, 8337975, Faks. (0251) 8338820
E-mail: imade.tasma@gmail.com, bb_biogen@litbang.pertanian.go.id

Diterima: 27 Maret 2015; Direvisi: 17 September 2015; Disetujui: 8 Oktober 2015

ABSTRAK

Sumber daya genetik (SDG) tanaman menyediakan materi dasar untuk program pemuliaan tanaman. Namun, baru sebagian kecil (<1%) koleksi SDG yang dimanfaatkan untuk pemuliaan tanaman. Karakterisasi SDG sudah banyak dilakukan dengan menggunakan karakter morfologi, namun metode ini lambat, menyita waktu, dan memerlukan banyak tenaga. Teknologi sekuensing modern menghasilkan peta genom rujukan suatu spesies tanaman yang dapat mempercepat karakterisasi SDG menggunakan teknik *next generation sequencing* (NGS). Tulisan ini mengulas pemanfaatan teknologi sekuensing genom untuk karakterisasi, proteksi, dan pemanfaatan SDG untuk mempercepat program pemuliaan tanaman. Di Indonesia, teknologi NGS telah dimanfaatkan sejak 2010 untuk resequencing genom komoditas unggulan nasional seperti kedelai, kakao, jagung, dan cabai merah. Jutaan SNP dan Indel telah diidentifikasi pada setiap komoditas sebagai sumber daya pemuliaan yang bernilai tinggi. Sebagian kecil SNP/Indel tersebut berada pada *protein coding region* yang potensial untuk penemuan gen-gen unggul. Selain SNP yang diidentifikasi pada semua genotipe, ditemukan SNP pada genotipe tertentu (SNP unik). Koleksi SNP dalam jumlah besar ini digunakan untuk mensintesis SNP chip untuk *genotyping* SDG secara cepat dan komprehensif. Didukung data fenotipe, SNP chip bermanfaat untuk melabel gen-gen unggul. Marka SNP yang berpautan dengan karakter unggul digunakan untuk menyeleksi individu pembawa karakter unggul tersebut. Dengan teknologi NGS, perakitan VUB tanaman dapat dilakukan lebih cepat, akurat, dan efisien. Dengan demikian, teknologi NGS dapat memfasilitasi karakterisasi dan pemanfaatan SDG untuk mempercepat program pemuliaan tanaman.

Kata kunci: Sumber daya genetik tanaman, sekuensing genom, keragaman genetik, pemuliaan tanaman

ABSTRACT

Plant genetic resources (PGRs) provide basic materials for cultivar development program. However, only a small portion (<1%) of PGR collection has been used for breeding purposes. PGR characterization has successfully been done using morphological characteristics. However, this method is slow, less accurate, and laborious. Modern sequencing techniques resulted reference genome sequences that accelerate PGR characterization through

genome resequencing method using NGS technology. This paper reviewed the usefulness of modern sequencing technology for PGR characterization to accelerate crop breeding programs. In Indonesia, the NGS has been used since 2010 for resequencing priority crops such as soybean, cocoa, maize, and chili pepper. Millions SNPs and Indels were identified from each crop species that are very important resources for breeding. Only small parts of the identified SNP/Indel were protein coding regions that are potentially used to discover superior genes. It was also discovered SNPs found in all genotypes or were unique to a particular genotype. The large number of SNP collections was used to synthesize high density SNP chip for comprehensive and quick PGR genotyping. Supported by valid phenotypic data, SNP chip was used for gene tagging of important traits and markers linked to the traits are used in assisting selection of individual plants having the traits. Using molecular breeding methods, cultivar development can be done in a more precise, efficient, and faster manners. Modern sequencing technology, therefore, facilitates a more comprehensive characterization and utilization of PGR collection to accelerate national breeding programs.

Keywords: *Plant genetic resources, genome sequencing, genetic diversity, plant breeding*

PENDAHULUAN

Sumber daya genetik (SDG) tanaman menyediakan materi dasar untuk perakitan varietas unggul baru (VUB). Secara global, lebih dari 7,4 juta aksesori berbagai spesies tanaman dikonservasi pada Bank Gen di berbagai negara yang jumlahnya sekitar 1.750 (Upadhyaya *et al.* 2013). Namun, baru sebagian kecil (<1%) dari koleksi SDG tersebut dimanfaatkan untuk program pemuliaan tanaman terutama karena kurangnya informasi (genotipe dan fenotipe) mengenai karakter yang bernilai ekonomi yang dimiliki oleh koleksi SDG tersebut.

Teknologi baru diperlukan untuk mengakselerasi pemanfaatan koleksi SDG pada Bank Gen untuk program pemuliaan tanaman. Hal ini dapat dilakukan dengan memperbaiki metode *genotyping* dan *phenotyping* serta meningkatkan akses untuk mendapatkan materi genetik yang disimpan pada Bank Gen di berbagai negara. Muncul

dan berkembangnya teknologi *next-generation sequencing* (NGS) serta menurunnya biaya sekuensing genom dapat mengatasi masalah dalam *genotyping* tersebut. Teknologi NGS dapat mempercepat pemetaan sekuen genom acuan (*reference genome sequence*) berbagai spesies tanaman, dari tanaman model sampai tanaman budi daya yang bernilai ekonomi tinggi, dari tanaman model *Arabidopsis thaliana* sampai tanaman yang berukuran genom besar dan kompleks seperti jagung, kelapa sawit, kedelai, tebu, dan gandum. Sampai saat ini, setidaknya 39 spesies tanaman telah tersedia peta genom acuannya (Dhanapal 2012; Van *et al.* 2013).

Peta genom acuan suatu spesies tanaman merupakan fondasi untuk penemuan gen-gen penting yang berkontribusi pada karakter-karakter bernilai ekonomi (Metzker 2010). Tersedianya peta genom acuan memungkinkan peneliti melakukan resekuensing untuk memahami secara lebih baik susunan genom individu koleksi SDG.

Informasi mengenai genotipe individu SDG dapat membantu mempercepat penemuan berbagai variasi genetik untuk mendukung program pemuliaan tanaman secara terarah, lebih cepat, dan lebih akurat dengan menggunakan metode baru, yaitu pemuliaan tanaman berbasis data genom. Program pemuliaan tanaman ke depan akan lebih fokus pada komparasi komposisi genom individu plasma nutfah. Hal ini akan membuka peluang penggunaan kombinasi strategi baru dalam pemetaan genetik dan analisis evolusi untuk penemuan dan pemanfaatan keragaman genetik SDG tanaman secara optimal untuk menunjang program pemuliaan yang bersinambungan.

Tulisan ini mengulas pemanfaatan teknologi sekuensing modern dalam karakterisasi, proteksi, dan pemanfaatan SDG secara efisien dan komprehensif untuk menunjang program pemuliaan tanaman.

TEKNOLOGI SEKUENSING DNA KAPASITAS TINGGI

Sekuensing DNA atau pengurutan basa DNA merupakan teknik kunci dalam perkembangan ilmu pengetahuan, di antaranya genetika, bioteknologi, biologi molekuler, dan genomika (Franca *et al.* 2002). Sekuensing DNA bertujuan untuk menentukan urutan basa nitrogen (adenin, guanin, sitosin, dan timin) suatu sampel DNA. Salah satu contoh aplikasi ambisius teknologi sekuensing DNA yaitu pengurutan genom manusia melalui proyek yang dikenal *Human Genome Project*. Dalam ilmu pengobatan, sekuensing DNA digunakan untuk mengidentifikasi, mendiagnosis, dan mengembangkan pengobatan penyakit genetik. Tahun 1970 merupakan awal pengembangan sekuensing DNA dengan metode kromatografi. Selanjutnya diperkenalkan metode sekuensing DNA dengan menggunakan metode *dye-based sequencing* (Olsvik *et al.* 1993).

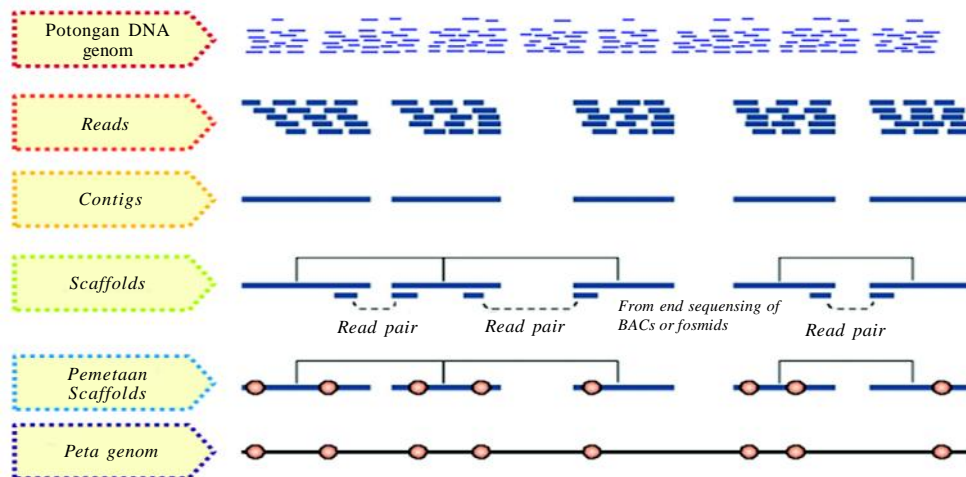
Pada awalnya, alat untuk membaca urutan basa DNA hanya mampu menentukan urutan basa DNA yang panjangnya kurang dari 1.000 basa. Oleh karena itu, untuk sekuensing genom, molekul DNA harus dipotong menjadi ukuran yang lebih kecil agar dapat dibaca oleh alat tersebut. Data urutan DNA yang dihasilkan kemudian disusun ulang menggunakan program komputer sehingga urutan keseluruhan DNA dalam inti sel dapat diketahui.

Sebelum ditemukannya alat yang mampu membaca urutan DNA secara paralel dalam jumlah besar, perurutan basa-basa DNA (*sequencing*) memakan waktu, tenaga, dan biaya yang sangat besar. Pemerintah Amerika Serikat, misalnya, menginvestasikan sekitar USD3,8 miliar (USD5,6 miliar jika memperhitungkan inflasi) untuk mencari urutan basa DNA manusia yang ukurannya sekitar tiga miliar pasang basa, yang baru dapat diselesaikan pada 2003 atau sekitar 13 tahun sejak proyek tersebut dicanangkan.

Setelah ditemukannya alat pembaca DNA paralel generasi berikutnya, waktu yang diperlukan untuk mendapatkan data yang sama dapat dipangkas menjadi hanya beberapa bulan. Namun, data yang diperoleh sangat besar sehingga memerlukan fasilitas komputer dan ahli bioinformatika untuk memrosesnya. Pusat-pusat penelitian genom terkemuka di dunia seperti Broad Institute, Washington University, maupun Beijing Genome Institute memiliki akses ke komputer berkapasitas besar serta ratusan ahli bioinformatika untuk mengelola dan memroses data sekuen DNA yang diperoleh.

Metode sekuensing DNA yang pertama kali digunakan adalah metode Sanger (*Sanger dideoxy sequencing*). Metode ini menggunakan DNA templat dan memerlukan primer spesifik untuk reaksi sekuensing. Teknologi sekuensing Sanger dan modifikasinya mendominasi metode sekuensing selama 30 tahun (Sanger *et al.* 1997). Panjang sekuen yang dihasilkan berkisar antara 1.000–1.200 pasang basa (bp) dan tidak mampu mencapai lebih dari 2.000 bp. Untuk mendapatkan sekuen yang lebih panjang dikembangkan metode sekuensing *shotgun* melalui proyek sekuensing genom manusia. Pada metode ini, templat DNA dipotong dengan enzim restriksi lalu fragmen DNA diklon pada vektor sekuensing dan individu fragmen DNA setiap klon disekuen terpisah. Hasil sekuen lengkap dari fragmen DNA yang panjang dapat diperoleh dengan menjajarkan (*alignment*) dan menyambung (*assembly*) sekuen fragmen DNA berdasarkan bagian sekuen yang berkomplemen (*overlapping*), yang memungkinkan pemetaan genom manusia dapat diselesaikan (Gambar 1).

Generasi kedua teknologi sekuensing setelah generasi pertama Sanger dikenal dengan *next generation sequencing* (NGS). Filosofi dasar pembentukan mesin NGS diadaptasi dari metode sekuensing *shotgun* (Venter *et al.* 2003; Shendure *et al.* 2004). Istilah NGS secara kolektif digunakan untuk mendeskripsikan semua teknologi sekuensing selain teknologi sekuensing Sanger. Teknologi ini berpotensi menekan biaya sekuensing satu



Gambar 1. Skema pembentukan peta sekuen genom acuan tanaman (Tasma *et al.* 2012). Gambar ini memperlihatkan proses pembacaan sekuen genom total dengan teknik *shotgun*. Analisis sekuen dimulai dari potongan kecil (100 bp) dibuat *contigs*, *scaffolds*, dan peta genom utuh. Proses ini memerlukan waktu panjang untuk mengisi celah (*gaps*) antar-*contigs* dan *scaffolds* yang memerlukan sekuensing dan analisis data sekuen berulang.

genom manusia hanya dengan USD1.000 (Kling 2005) dan teknologi semacam ini sudah cukup lama tersedia di pasar.

Teknologi NGS membaca templat DNA secara acak (*random*) sepanjang seluruh genom dengan membuat potongan-potongan pendek DNA genom, kemudian menyambungkannya dengan *adapter* (potongan DNA pendek yang didesain khusus untuk tujuan ini) agar dapat dibaca oleh mesin NGS secara random selama proses sintesis DNA. Dengan demikian, teknologi NGS sering disebut dengan sekuensing paralel secara masif (*massively parallel sequencing*). Panjang bacaan sekuen DNA yang dihasilkan mesin NGS jauh lebih pendek dibandingkan bila menggunakan mesin sekuensing dengan metode Sanger. NGS menghasilkan panjang sekuen DNA antara 50–500 bp. Karena sekuen yang dihasilkan NGS pendek, sekuensing setiap fragmen DNA mesti dilakukan lebih dari sekali ukuran genom (*genome sequence coverage*). Sebagai contoh, bila *sequence coverage* suatu genom 30 kali berarti fragmen-fragmen DNA pada genom tersebut disekuensi 30 kali sehingga setiap fragmen DNA dalam genom dibaca mesin 30 kali. Hal ini dilakukan untuk menjaga akurasi data hasil sekuensing.

Penerapan inovasi teknologi NGS berkembang sangat cepat dan berbagai kemajuan telah dicapai. Tiga teknologi NGS utama yang tersedia saat ini adalah *Roche/454 pyrosequencing platform*, *Illumina Solexa polymerase sequencing platform*, dan *ABI/SOLID ligase sequencing technology*. Dibandingkan dengan metode sekuensing Sanger, ketiga teknologi NGS ini menghasilkan data sekuen yang jauh lebih banyak dalam sekali menjalankan alat. Dengan demikian, alat ini dikenal dengan *high throughput sequencing platforms* (Ansorge 2009). Prinsip dasar ketiga alat NGS tersebut berbeda, baik dalam menghasilkan data sekuen, kualitas data yang

dihasilkan maupun biaya sekuensing (Tabel 1). Alat Roche/454 menghasilkan data sekuen terpanjang, tetapi kuantitas data sekuennya terendah (*lowest throughput*). Jumlah sekuen basa yang dihasilkan Illumina/Solexa tertinggi (*highest throughput*), tetapi panjang bacaannya hanya sekitar 100 basa. Panjang bacaan yang dihasilkan ABI/SOLID terpendek, hanya sekitar 50 basa, tetapi tingkat kesalahan pembacaan DNA pada proses sekuensing (*error rate*) paling rendah.

Teknologi NGS merupakan revolusi metode sekuensing dan menghasilkan data sekuen yang luar biasa besar dalam waktu relatif singkat dibandingkan dengan metode sekuensing Sanger. Oleh karena itu, diperlukan sumber daya teknologi informasi dan ahli bioinformatika untuk menganalisis data yang dihasilkan sehingga bermanfaat untuk mendukung program pemuliaan tanaman bernilai ekonomi tinggi. Badan Litbang Pertanian saat ini memiliki NGS HiSeq2000 yang terutama dimanfaatkan untuk penelitian resekuensing komoditas unggulan Kementerian Pertanian seperti jagung, kedelai, sapi, kakao, sawit, cabai merah, kentang, pisang, dan jarak pagar.

Saat ini telah diperkenalkan alat sekuensing generasi ketiga (*third generation sequencing platforms*) yang mampu melakukan sekuensing molekul DNA atau RNA, yaitu *HeliScope*, *Ion Torrent*, *Single Molecule Real-Time Sequencing (SMRT)*, dan *Oxford Nanopore*. Keunggulan alat-alat tersebut adalah mampu membaca urutan molekul DNA yang berasal dari molekul tunggal. Dengan demikian, kebutuhan materi awal untuk sekuensing sangat sedikit (< 1 µg), baik DNA maupun RNA, dan tidak diperlukan amplifikasi PCR sehingga dapat terhindar dari kesalahan (*bias*) yang disebabkan oleh amplifikasi PCR. Namun, data sekuen yang dihasilkan (*sequence throughput*) dari alat sekuensing generasi

Tabel 1. Perbandingan teknologi alat sekuensing generasi kedua.

Nama alat	Instrumen	Metode sekuensing	Lama menjalankan alat (hari)	Jumlah basa per pembacaan (bp)	Hasil sekuensing sekali menjalankan alat	Tingkat kesalahan pembacaan (%)	Keuntungan alat	Kerugian alat
454	454 FLX+	<i>Pyrosequencing</i>	18–20	700	900	1	Hasil pembacaan panjang	Biaya tinggi untuk tiap Mega basa (Mb)
Illumina	Illumina HiSeq2000-V3	<i>Synthesis</i>	10	100+100	600.000	0,1	Hasil pembacaan terbanyak (<i>highest throughput</i>), biaya per Mb termurah	Biaya modal tinggi dan memerlukan komputasi tinggi
SOLID	SOLID 5500 xl (4hq)	Ligation	8	75 + 35	155.100	0,01	Hasil pembacaan tinggi (<i>high throughput</i>)	Hasil bacaan pendek, lebih banyak bolong saat analisis penyambungan dan biaya per Mb lebih mahal

Sumber: Gao *et al.* (2012).

ketiga ini lebih rendah dibandingkan yang dihasilkan alat sekuensing generasi kedua (NGS), selain tingkat kesalahannya masih tinggi. Dengan keterbatasan tersebut, alat sekuensing generasi ketiga ini hanya dapat digunakan untuk sekuensing organisme dengan ukuran genom kecil seperti mikroba. Dengan demikian, dari aspek kuantitas data sekuen yang dihasilkan (*throughput*) dan biaya sekuensing, alat sekuensing generasi ketiga ini belum dapat menandingi kemampuan alat sekuensing generasi kedua (NGS).

Aplikasi teknologi NGS untuk resekuensing genom SDG tanaman akan sangat bermanfaat dalam mempercepat program pemuliaan tanaman. Resekuensing membutuhkan peta genom acuan, sedangkan sekuensing *de novo* (*de novo assembly*) tidak memerlukannya. Sekuensing *de novo* genom tanaman menggunakan NGS belum memberikan hasil yang memuaskan karena tingginya tingkat kompleksitas genom tanaman akibat proses duplikasi genom dan pengayaan genom tanaman melalui sekuen berulang (*repeated sequences*) (Varshney *et al.* 2009). Oleh karena itu, teknologi NGS lebih tepat digunakan untuk resekuensing genom spesies tanaman yang telah memiliki peta sekuen genom acuan, terutama untuk identifikasi *single nucleotide polymorphisms* (SNPs) sebagai marka DNA, mendapatkan alel fungsional baru, dan menguji pola seleksi selama proses domestikasi tanaman.

Genotyping genom total dengan marka SNP (*genome-wide SNP genotyping*) sangat penting dalam studi pemetaan asosiasi dan evolusi genom (Akhunov *et al.* 2009). Panel marka SNP yang dikembangkan dengan konsorsium berbagai spesies tanaman, misalnya padi dan

apel, aplikasinya tidak mencakup kelompok SDG secara luas. Teknologi *genotyping by sequencing* (GBS) dapat memecahkan masalah ini dengan menyediakan lebih banyak marka yang polimorfis (Huang *et al.* 2010). Teknologi GBS mengembangkan marka berdasar genom total seperti *massively developed restriction site-associated DNA* (RAD) yang digunakan untuk mengkonstruksi peta genetik densitas tinggi dalam analisis keragaman genetik tanaman (Gupta *et al.* 2008).

PEMBENTUKAN PETA SEKUEN GENOM ACUAN

Langkah awal dalam mengkarakterisasi individu plasma nutfah pada level genom adalah pembuatan peta genom acuan suatu spesies tanaman. Pembentukan peta genom acuan perlu dilakukan jika spesies tanaman yang dipelajari belum memiliki informasi DNA yang lengkap. Mengingat sifat-sifat makhluk hidup dikode oleh DNA, pengetahuan tentang komposisi DNA organisme yang dipelajari sangat diperlukan untuk memahami dan memanfaatkan sifat-sifat unggul yang dimiliki organisme tersebut. Penentuan komposisi dan urutan basa-basa DNA merupakan kegiatan yang rumit mengingat tanaman memiliki paling tidak 100 juta hingga miliaran pasang basa dalam inti selnya. Basa-basa DNA ini memiliki ukuran sub-mikroskopis dan mudah terdegradasi apabila tidak ditangani dan disimpan secara benar.

Sekuen genom acuan setiap spesies tanaman target diperlukan oleh peneliti genomika untuk memanfaatkan

secara maksimal kemajuan dan keunggulan teknologi genomik untuk penelitian pertanian. Untuk menghasilkan dan menganotasi sekuen genom suatu spesies tanaman pertama kali, DNA kromosom dipotong-potong menjadi miliaran potongan DNA dengan panjang antara 500–30.000 bp, dan berbagai jenis set pustaka genom (*complex DNA libraries*) dihasilkan dalam beberapa bulan. *High cost long read sequence data* diproduksi dari *complex DNA libraries* dan pemetaan *long contiguous chromosomal regions* dikonstruksi menggunakan *de novo assembly computer*. Sekuen yang panjang memungkinkan program komputer untuk menyambung (*assembly*) puluhan ribu sekuen secara berulang (*repetitive element*) yang umumnya terdapat pada genom tanaman, yang pada akhirnya menghasilkan *sequence scaffolds* yang lebih panjang (Gambar 1). *Sequence scaffolds* yang lebih panjang akan lebih bermanfaat bagi peneliti genomika dibandingkan dengan yang ukurannya lebih pendek. Sekali suatu *sequence scaffold* panjang diperoleh, *express regions* dari genom dianotasikan dengan jutaan sekuen pustaka cDNA yang juga diperoleh menggunakan *long read sequencing technologies*. Sangat mirip dengan sekuensing genom *de novo*, *long read sequencing technologies* untuk sekuensing cDNA lebih disukai daripada *NGS short read sequencing technologies* untuk konstruksi draf pertama *transcriptome*. Hal ini karena sekuen yang lebih panjang ketika disambung (*assembly*) akan lebih besar kemungkinannya menghasilkan sekuen transkrip utuh yang lebih bermanfaat untuk gen model (Gambar 1).

Sekuen genom acuan berkualitas tinggi (panjang mendekati ukuran genom sesungguhnya) dan akurat (tingkat kesalahan rendah) jauh lebih mahal memproduksinya dibandingkan dengan membuat data resequencing genom yang sama. Tim peneliti tidak pernah dapat menyelesaikan sekuen acuan, tetapi akan menyempurnakan sekuen genom acuan yang telah dipublikasikan sebelumnya. Sebagai contoh, walaupun sekuen genom acuan manusia pertama kali diselesaikan dan dipublikasi pada tahun 2002, sekitar 300 peneliti terus mengedit sekuen acuan tersebut dan sekarang telah diperoleh sekuen genom acuan generasi ke-19. Makin akurat sekuen genom acuan suatu spesies tanaman, makin akurat dan makin kuat data resequencing yang dihasilkan.

Untuk sekuensing suatu spesies tanaman, peneliti kadang-kadang perlu membuat 3–5 sekuen genom acuan dari beberapa spesies yang berkerabat sangat dekat dengan spesies tanaman yang sedang dipelajari. Sebagai contoh, pada jagung, sekuen genom acuan sudah dibuat dari B73, MO17, dan spesies *popcorn* (Calzada *et al.* 2009; Schnable *et al.* 2009). Pada padi, sekuen genom acuan dibuat dari Niponbare, Indica, dan Javanica (Goff *et al.* 2002; Yu *et al.* 2002). Pada kakao, sekuen genom acuan dibuat dari Belizean Criollo dan dari kultivar Matina 1–6 (Argout *et al.* 2011). Pada kelapa sawit (*Elaeis guineensis*), sekuen genom acuan dibuat dari genotipe Dura, Pisifera, dan spesies *E. oleifera* (Singh *et al.* 2013).

Salah satu aplikasi teknologi sekuensing ialah karakterisasi genom spesies tanaman untuk membuat sekuen genom acuan suatu spesies tanaman. Sekuen genom acuan menjadi pedoman dalam membuat peta genom untuk studi lanjutan berbagai individu aksesi anggota spesies tanaman tersebut.

Pembentukan peta genom acuan tanaman tidak dapat dilepaskan dari pemanfaatan teknologi NGS yang berkembang demikian cepat. Revolusi teknologi sekuensing DNA dengan teknologi NGS membuat biaya sekuensing genom menjadi murah sehingga memungkinkan melakukan sekuensing lebih banyak genom spesies tanaman dengan biaya yang terjangkau (Varshney *et al.* 2009). Sekitar 11 spesies tanaman telah tersedia peta genom acuannya yang diselesaikan menggunakan teknologi NGS (Tabel 2). Spesies tanaman yang peta genom sekuennya diselesaikan menggunakan teknologi NGS yaitu mentimun (*Cucumis sativus*), apel (*Malus domestica*), stroberi hutan (*Fragaria vesca*), kakao (*Theobroma cacao*), kedelai liar (*Glycine soja*), jarak pagar (*Jatropha curcas*), *date palm* (*Phoenix dactylifera*), kentang (*Solanum tuberosum*), *thellungella* (*Thellungella parvula*), sawi China (*Brassica rapa*), dan *Indian hemp* (*Canabis sativa*).

Genom *Arabidopsis thaliana* selesai disekuen pada Desember 2000, merupakan sekuen genom acuan tanaman pertama yang dipublikasi dan dapat diakses oleh publik (Cao *et al.* 2011). Berikutnya puluhan spesies tanaman telah selesai disekuen genom acuannya (Tabel 1). Peranan teknologi NGS sangat nyata dalam mengakselerasi pembentukan sekuen genom acuan berbagai spesies tanaman (Tabel 2). Saat ini lebih dari 42 peta sekuen genom acuan tanaman budi daya yang meliputi 39 spesies telah tersedia untuk publik dan peta genom sekuen dari 12 spesies sedang diselesaikan atau belum dapat diakses publik (Van *et al.* 2013). Peta genom acuan 39 spesies tanaman tersebut kebanyakan diselesaikan setelah tahun 2005 ketika teknologi NGS dikembangkan.

Tersedianya peta sekuen genom acuan berbagai spesies tanaman budi daya mengakselerasi penemuan gen maupun marka molekuler (SNP dan Indel) yang berasosiasi dengan karakter bernilai ekonomi tinggi. Hal ini pada akhirnya mempercepat program pemuliaan tanaman untuk menghasilkan kultivar-kultivar unggul baru dalam rangka mendukung ketahanan pangan nasional yang berkelanjutan. Tersedianya peta sekuen genom acuan spesies tanaman akan mengakselerasi studi evolusi tanaman secara lebih detail, dari level gen ke level nukleotida. Hal ini akan sangat bermanfaat dalam memahami kompleksitas genom dan hubungan genotipe dan data evolusi tanaman. Karena duplikasi genom dan variasi struktur kromosom berperan sangat penting dalam proses evolusi tanaman, perkembangan teknologi NGS memungkinkan penemuan gen-gen baru dengan fungsi baru melalui investigasi keragaman fungsional dan keragaman evolusi berbagai spesies tanaman yang telah tersedia peta sekuen genom acuannya.

Tabel 2. Berbagai spesies tanaman yang telah mempunyai peta sekuen genom acuan.

Spesies tanaman	Nama latin	Ukuran genom (Mbp)*	Referensi
Arabidopsis	<i>Arabidopsis thaliana</i>	207	The Arabidopsis Thaliana Genome Initiative (2000); Cao <i>et al.</i> (2011)
	<i>Arabidopsis lyrata</i>	207	Hu <i>et al.</i> (2011)
Padi	<i>Oryza sativa ssp indica</i>	430	Yu <i>et al.</i> (2002)
	<i>O. sativa ssp japonica</i>	430	Goff <i>et al.</i> (2002)
Poplar	<i>Populus trichocarpa</i>	485	Tuskan <i>et al.</i> (2006)
Anggur	<i>Vitis vinifera</i>	487	Jallion <i>et al.</i> (2007)
Teratai (Lotus)	<i>Lotus japonicas</i>	472	Sato <i>et al.</i> (2008)
Pepaya	<i>Carica papaya</i>	372	Ming <i>et al.</i> (2008)
Jagung	<i>Zea mays</i>	2.300	Schnable <i>et al.</i> (2009)
	<i>Zea mays ssp. parviglumis</i>	2.100	Calzada <i>et al.</i> (2009)
Kapas	<i>Gossypium raimonddi</i>	750	www.phytozome.net/cotton.php
Sorgum	<i>Sorghum bicolor</i>	730	Paterson <i>et al.</i> (2009)
Mentimun	<i>Cucumis sativus</i>	367	Huang <i>et al.</i> (2009)
	<i>Cucumis sativus</i>	368	Wóycicki <i>et al.</i> (2011)
Kentang	<i>Solanum tuberosum</i>	844	Xu <i>et al.</i> (2011)
Kanola	<i>Brassica napus</i>	302	Wang <i>et al.</i> (2011)
Pisang	<i>Musa accuminata</i>	472,2	D'Hont <i>et al.</i> (2012)
Kakao	<i>Theobroma cacao</i>	430	Argout <i>et al.</i> (2011)
Jarak kepyar	<i>Ricinus communis</i>	350	Chan <i>et al.</i> (2011)
Apel	<i>Malus domestica</i>	742,3	Velasco <i>et al.</i> (2010)
Stroberi	<i>Fragaria vesca</i>	240	Shulaev <i>et al.</i> (2011)
Kedelai	<i>Glycine max</i>	1.115	Schmutz <i>et al.</i> (2010)
Kacang gude	<i>Cajanus cajan</i>	833,07	Varshney <i>et al.</i> (2012)
Alfalfa	<i>Medicago sativa</i>		Young <i>et al.</i> (2011)
Date palm	<i>Phoenix dactylifera</i>	658	Al-Dous <i>et al.</i> (2011)
Linum (<i>Flax</i>)	<i>Linum usitatissimum</i>	350	http://www.phytozome.net/flax
Semangka	<i>Citrullus lanatus</i>	450	Rent <i>et al.</i> (2012)
Kayu putih	<i>Eucalyptus grandis</i>	691	http://www.phytazome.net/eucaliptus.php
Foxtail millet	<i>Setaria italica</i>	510	Bennetzen <i>et al.</i> (2012)
Melon	<i>Cucumis melo L.</i>	430	Garcia-Mas <i>et al.</i> (2012)
Kelapa sawit	<i>Elaeis guineensis</i>	1.800	Singh <i>et al.</i> (2013)
Ubi kayu	<i>Manihot esculenta</i>	760	http://www.phytozome.net/cassava.php
Tomat	<i>Solanum lycopersicum</i>	900	The Tomato Genome Consortium (2012)
Buncis	<i>Phaseolus vulgaris</i>	486,9	http://www.phytozome.org/com/monbean.php

Sumber: Dhanapal (2012); Van *et al.* (2013).

RESEKUENSING GENOM UNTUK KARAKTERISASI VARIASI GENOM INDIVIDU

Penelitian resequencing biasanya dilakukan untuk mengetahui perbedaan/variasi DNA individu SDG tanaman dibandingkan dengan sekuen genom acuan yang ada. DNA individu yang dipelajari (misalnya DNA lima varietas padi nasional) diekstraksi dan dibaca urutannya dengan mesin NGS kemudian dibandingkan dengan sekuen genom acuan menggunakan analisis bioinformatika. Data variasi DNA yang diperoleh dapat dimanfaatkan untuk membedakan individu-individu dalam spesies yang sama (misalnya antara padi Rojolele dan IR64) serta mencari basis genetik sifat-sifat unggul yang dimiliki individu tertentu (misalnya sekuen DNA yang menyebabkan nasi dari padi Rojolele lebih wangi daripada IR64).

Data variasi di tingkat DNA sangat bermanfaat untuk berbagai macam keperluan, mulai dari perlindungan varietas hingga pemuliaan, karena DNA dapat diambil dari berbagai organ individu yang dipelajari serta tidak dipengaruhi oleh faktor lingkungan. Dengan demikian, identifikasi suatu individu maupun sifat-sifat unggul yang dibawanya dapat dilakukan kapan saja sepanjang sampel DNA dapat diperoleh dari individu tersebut.

Sekali suatu peta genetik sekuen genom acuan suatu spesies tanaman tersedia, peneliti genomika dapat melakukan resequencing individu aksesori kedua, ketiga, keempat, kelima, dan seterusnya dari anggota spesies yang sama menggunakan *very low-cost short read technologies* (NGS) seperti HiSeq2000 dari Illumina. Untaian sekuen pendek (100–150 bp) tidak perlu di-assembly satu dengan yang lainnya, tetapi dipetakan kembali ke sekuen genom acuan untuk mengidentifikasi perbedaan genetik antara sekuen acuan dan sekuen individu kedua, ketiga, dan seterusnya. Dengan demikian,

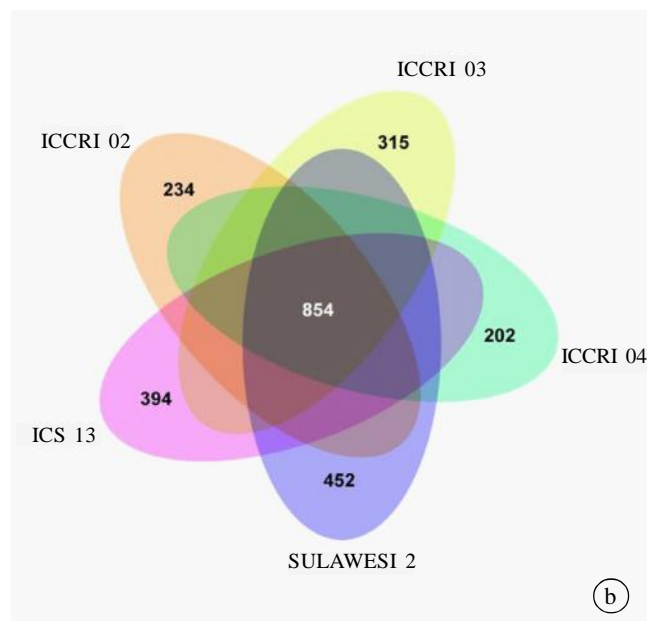
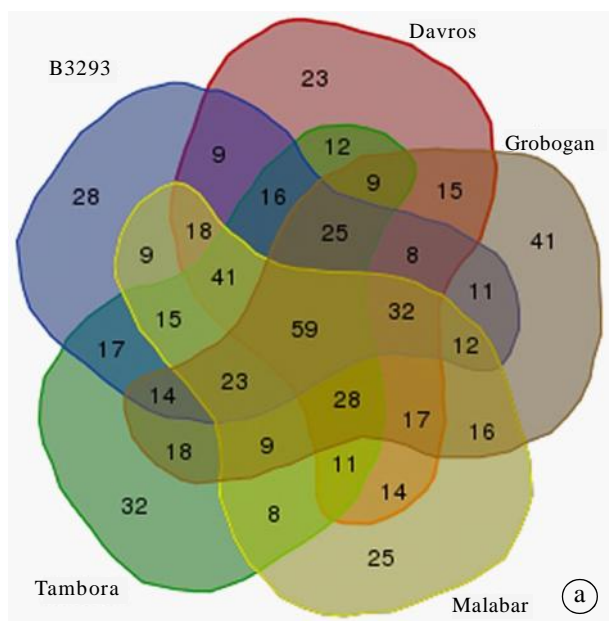
sejumlah perbedaan dapat didokumentasikan dan perbedaan genetik ini digunakan untuk menjelaskan perbedaan antargenotipe (varietas) dalam suatu spesies, misalnya daya hasil, ketahanan hama dan penyakit, waktu pembungaan, serta toleransi kekeringan dan banjir (Long dan Ort 2010). Penelitian resequencing dapat digunakan dalam survei keragaman dalam spesies, *genotyping* dengan sekuensing genom total, konstruksi peta genetik resolusi tinggi, dan pembuatan *genome wide association studies* (GWAS).

Resequencing berbagai galur unggul (*superior lines*) dan individu aksesori plasma nutfah dapat mengidentifikasi keunikan setiap genotipe SDG tanaman. Hal ini akan menjadi identitas dari varietas, galur, maupun individu aksesori sebagai sidik jari dari individu tersebut. Contoh dari keunikan ini disajikan pada Gambar 2 (Satyawana *et al.* 2014; Tasma *et al.* 2015a).

Sekuensing lima varietas kedelai Indonesia menghasilkan sekitar 2,6 juta variasi genom (SNP dan Indel). Sebagian besar SNP/Indel dideteksi di luar gen. Hanya sebagian kecil SNP/Indel dideteksi dalam gen. Hanya 2,15% variasi genom yang terdeteksi pada penelitian ini terdapat pada *protein coding region* (ekson) (Satyawana *et al.* 2014). Dari 337 SNP yang ada dalam ekson, ditemukan 59 SNP yang dapat membedakan kelima varietas kedelai Indonesia dari varietas rujukan Williams 82. Beberapa alel SNP hanya ditemukan dalam satu varietas dan berpotensi sebagai sidik jari varietas tersebut. Ditemukan SNP umum dan SNP unik pada lima genotipe kedelai. Genotipe B3293, Davros, Grobogan, Malabar, dan Tambora memiliki genotipe yang sama di 59 lokus SNP, yang berbeda dengan Williams 82. Pada setiap genotipe terdapat 10–16 SNP unik yang dapat berfungsi sebagai sidik jari (Gambar 2).

Hasil yang mirip ditemukan dalam resequencing lima varietas kakao Indonesia. Kelima klon kakao yang diresequen memiliki variasi DNA yang sama pada 854 lokus, yang berbeda dengan DNA genom rujukan. Ditemukan variasi DNA yang unik pada satu varietas, 234 lokus SNP untuk ICCRI 02, 315 lokus untuk ICCRI 03, 202 lokus untuk ICCRI 04, 452 lokus untuk Sulawesi 2, dan 394 lokus SNP untuk ICS 13 (Tasma *et al.* 2015a). SNP unik pada satu varietas dapat digunakan untuk karakterisasi dan sebagai sidik jari untuk memproteksi varietas tersebut. SNP unik yang berada pada setiap genotipe kedelai atau kakao berfungsi sebagai identitas unik atau sidik jari setiap genotipe tersebut.

Variasi SNP juga dijumpai pada komoditas lain seperti padi, kelapa sawit, pisang, dan kentang. SNP yang berfungsi sebagai sidik jari ini digunakan untuk memproteksi materi genetik dan bahan tanaman unggul dari pemanfaatan oleh orang yang tidak bertanggung jawab. Dengan sidik jari ini, peneliti dapat mendeteksi keberadaan materi genetiknya secara cepat dengan melakukan tes untuk melihat pola pita DNA SNP dari materi genetik miliknya. Secara lebih luas, teknologi sidik jari dapat digunakan untuk memproteksi materi genetik koleksi SDG nasional dari pemanfaatan yang tidak bertanggung jawab. Ini penting untuk menjaga kekayaan SDG pertanian nasional dan sebagai perlindungan hak kekayaan intelektual SDG tanaman unggul nasional. Dengan demikian, teknologi sekuensing modern menghasilkan sidik jari berdasar SNP yang deteksinya dapat dilakukan secara cepat, tepat, dan akurat untuk perlindungan kekayaan SDG nasional.



Gambar 2. SNP umum dan SNP unik beberapa varietas kedelai (a) dan SNP umum dan SNP unik beberapa varietas kakao (b) (Satyawana *et al.* (2014); Tasma *et al.* (2015a).

TEKNOLOGI NGS UNTUK AKSELERASI PEMANFAATAN SDG TANAMAN

Pemuliaan molekuler merupakan suatu metode pemuliaan tanaman dengan memanfaatkan teknologi genetika konvensional, biologi molekuler, dan genetika molekuler sebagai landasan teori untuk menggabungkan teknologi biologi modern ke dalam metode pemuliaan konvensional, sementara dilakukan penyaringan kombinasi genotipe/fenotipe dan galur-galur pemuliaan unggul untuk merakit varietas tanaman yang beradaptasi lebih baik. Tersedianya teknologi NGS dan peta genom total acuan spesies tanaman seperti padi, jagung, kedelai, kakao, kelapa sawit, pisang, dan kentang membuka peluang untuk mengeksplorasi keragaman pada level DNA di antara individu tanaman dalam suatu spesies dan hubungannya dengan keragaman fenotipe (Paterson *et al.* 2010). Tujuan akhir dari resequencing adalah mengetahui dasar molekuler hubungan data fenotipe dan data genotipe.

Data panel keragaman ratusan sampai ribuan genotipe tanaman dan peta genom acuan spesies tersebut yang dihasilkan dengan menggunakan teknologi NGS, merupakan alat untuk mengetahui keragaman genetik, asosiasi gen dengan fenotipe, dan eksploitasi keragaman genetik alami untuk pengembangan genotipe unggul. Untuk dapat melakukan hal ini secara efektif, diperlukan koleksi data fenotipe secara ekstensif dari individu anggota spesies untuk membuat panel keragaman dikombinasikan dengan data resequencing.

Mengoleksi data fenotipe ini akan menjadi tantangan berat dalam pemanfaatan data genomika dalam program pemuliaan modern. Alasan utamanya adalah karakter fenotipe yang bernilai ekonomi tinggi (misal daya hasil) umumnya bersifat kuantitatif yang dikendalikan oleh banyak gen dan memerlukan pengamat yang sangat berpengalaman untuk menyekor data fenotipe secara efektif dan konsisten. Kondisi ini menyebabkan kemajuan dalam koleksi data fenotipe belum berjalan seiring dengan koleksi data genotipe menggunakan metode genomika NGS. Di pihak lain, peneliti/teknisi yang terlatih dan berdedikasi tinggi untuk mengoleksi data fenotipe sesuai bidang keahliannya jumlahnya sedikit. Ke depan diperlukan paradigma baru untuk melatih peneliti dalam bidang pemuliaan tanaman. Pemulia tanaman perlu mengaplikasikan teknik genomika modern seperti sekuensing dan mengembangkan metode baru untuk koleksi data fenotipe (*phenomics*) dalam upaya memanfaatkan kemajuan teknologi genomika untuk perakitan varietas baru secara efisien, efektif, dan sinambung.

Berkembangnya teknologi sekuensing modern dan biaya sekuensing yang semakin terjangkau mengubah paradigma penanganan, karakterisasi, proteksi, dan pemanfaatan SDG tanaman dari semula berdasarkan data morfologi tanaman menjadi berdasarkan data genomika. Sekuensing genom dapat menganalisis genotipe SDG

tanaman secara cepat, tepat, akurat, dan komprehensif sehingga pemanfaatannya untuk program pemuliaan tanaman juga lebih komprehensif dan berpresisi tinggi. Teknologi sekuensing genom juga telah merevolusi teknologi pemuliaan tanaman dari seleksi berdasarkan data fenotipe menjadi seleksi berdasarkan data genom yang jauh lebih efisien, efektif, komprehensif, dan berpresisi tinggi.

Di Indonesia, pemuliaan tanaman berbasis data genom kapasitas tinggi ini telah mulai dilakukan sejak tahun 2010 dengan membentuk konsorsium penelitian pemuliaan berbasis data genom. Anggota konsorsium adalah Balai-Balai Penelitian komoditas di lingkup Badan Litbang Pertanian. Penelitian dimulai dengan melakukan resequencing anggota spesies tanaman unggulan nasional seperti kedelai, jagung, kelapa sawit, kakao, pisang, cabai merah, dan kentang. Variasi SNP dan Indel yang berhasil diidentifikasi melalui penelitian ini berkisar antara 2,69 juta SNP/Indel pada kakao hingga 5,73 juta SNP/Indel pada pisang (Tasma *et al.* 2015b). Di antara variasi SNP/Indel yang terdeteksi, hanya sebagian kecil yang ada pada ekson (*protein coding region*), berkisar antara 0,86% pada kakao sampai 8,03% pada kentang. Variasi SNP/Indel yang diidentifikasi umumnya berada di luar gen.

Variasi SNP/Indel ini sebagian kecil sudah diverifikasi menggunakan metode sekuensing Sanger dan amplifikasi PCR menggunakan teknik molekuler sederhana. SNP/Indel yang telah diverifikasi digunakan untuk mensintesis SNP chip kapasitas tinggi. Saat ini sedang didesain SNP chip kedelai yang mengandung sedikitnya 3.000 marka SNP. SNP chip ini dapat digunakan untuk pelabelan gen-gen unggul bernilai ekonomi tinggi untuk mendukung program pemuliaan kedelai nasional. Sintesis SNP chip seperti yang dilakukan pada kedelai akan dilakukan pula pada kakao, kelapa sawit, cabai merah, dan pisang untuk mendukung percepatan pelabelan gen-gen unggul dan pemuliaan karakter-karakter penting pada setiap komoditas.

Untuk pemuliaan jangka pendek, beberapa karakter target pada setiap komoditas tanaman yang masuk dalam kegiatan konsorsium adalah toleransi kekeringan pada jagung, toleransi keracunan aluminium dan komponen hasil tinggi pada kedelai, ketahanan penyakit busuk buah (*Phytophthora palmivora*) pada kakao, *slow stem growth rate* dan kadar minyak tinggi pada kelapa sawit, ketahanan penyakit antraknose pada cabai merah, penyakit virus (PVX, PVY, PLRV) pada kentang, dan ketahanan penyakit layu Fusarium pada pisang (Tasma *et al.* 2015b).

Marka SNP/Indel hasil penelitian resequencing digunakan untuk melabel gen-gen pengkode karakter-karakter penting pada setiap komoditas yang diteliti dengan penelitian menggunakan populasi yang dikembangkan Balai-Balai. Marka DNA yang berpautan dengan karakter target digunakan untuk pemuliaan berbasis genom untuk mempercepat perolehan varietas unggul baru dengan karakter yang diinginkan.

KESIMPULAN

Teknologi sekuensing berkembang pesat dan menghasilkan alat berkapasitas tinggi NGS. Alat tersebut telah mengakselerasi penyelesaian peta genom acuan berbagai spesies tanaman.

Peta genom acuan berfungsi sebagai referensi dalam analisis variasi genom individu SDG menggunakan teknologi resequencing. Penjajaran data sekuen dengan sekuen genom acuan menghasilkan SNP dan Indel dalam jumlah besar sebagai sumber daya marka DNA untuk mendukung program pemuliaan tanaman. Sebagian kecil variasi tersebut adalah SNP/Indel yang berlokasi pada ekson yang dapat berujung pada isolasi gen unggul. Resequencing juga menghasilkan SNP unik yang berfungsi sebagai sidik jari untuk proteksi SDG.

Koleksi SNP digunakan untuk sintesis SNP *chip* kapasitas tinggi untuk mempercepat pelabelan karakter unggul. SNP yang berpautan dengan karakter unggul dapat diaplikasikan dalam program pemuliaan berbantuan marka.

Penelitian sekuensing genom di Indonesia telah dikembangkan pada komoditas unggulan pertanian untuk menemukan gen-gen unggul untuk mempercepat program pemuliaan tanaman. Melalui teknologi ini telah diperoleh SNP umum dan SNP unik yang bermanfaat untuk pengembangan SNP *chip* kapasitas tinggi dan sebagai sidik jari untuk proteksi SDG nasional.

Ke depan, dengan semakin murahnya biaya sekuensing genom, karakterisasi aksesori SDG dengan sekuensing genom semakin terjangkau dan pemanfaatan koleksi SDG untuk program pemuliaan lebih intensif dan komprehensif. Tantangannya ialah perlunya kemampuan pemulia menyediakan data fenotipe aksesori SDG mengingat perkembangan teknologi *phenotyping* belum secepat teknologi sekuensing genom.

DAFTAR PUSTAKA

- Akhunov, E., C. Nicolet, and J. Dvorak. 2009. Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina Goldengate assay. *Theor. Appl. Genet.* 119: 507–517.
- Al-Dous, E.K., B. George, M.E. Al-Mahmoud, M.Y. Al-Jaber, H. Wang, Y.M. Salameh, E.K. Al-Azwani, S. Chaluvadi, A.C. Pontaroli, J. DeBarry, V. Arondel, J. Ohlrogge, I.J. Saie, K.M. Suliman-Elmeer, J.K. Bennetzen, R.R. Kruegger, and J.A. Malek. 2011. *De novo* genome sequencing and comparative genomics of date palm (*Phoenix dactylifera*). *Nat. Biotechnol.* 29: 521–527.
- Ansorge, W.J. 2009. Next generation DNA sequencing techniques. *Nat. Biotechnol.* 25: 195–203.
- Arabidopsis Thaliana Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nat. Biotechnol.* 408: 796–815.
- Argout, X. *et al.* [68 authors]. 2011. The genome of *Theobroma cacao*. *Nat. Genet.* 43: 101–108.
- Bennetzen, J.L. *et al.* [34 authors]. 2012. Reference genome sequence of the model plant *Setaria*. *Nat. Biotechnol.* 30: 555–561.
- Calzada, J.P.V., O.M. de la Vega, G. Hernandez-Guzman, E. Ibarra-Laclette, C. Alvarez-Mejia, J.C. Vega-Arreguin, B. Jimenez-Moraila, A. Fernandez-Cortes, G. Corona-Armenta, L. Herrera-Estrella, and A. Herrera-Estrella. 2009. The Palomero genome suggests metal effects on domestication. *Science* 326: 1078.
- Cao, J., K. Schneeberger, S. Ossowski, T. Günther, S. Bender, J. Fitz, D. Koenig, C. Lanz, O. Stegle, C. Lippert, X. Wang, F. Ott, J. Müller, C. Alonso-Blanco, K. Borg-Wardt, K.J. Schmid, and D. Weigel. 2011. Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. *Nat. Genet.* 43: 956–965.
- Chan, A.P. *et al.* [18 authors]. 2011. Draft genome sequence of the oilseed species *Ricinus communis*. *Nat. Biotechnol.* 28: 951–956.
- Dhanapal, A.P. 2012. Genomics of crop plant genetic resources. *Adv. Biosci. Biotechnol.* 3: 378–385.
- D’Hont, A. *et al.* [62 authors]. 2012. The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature* 488: 213–219.
- Franca, L.T.C., E. Carrilho, and T.B.L. Kist. 2002. A review of DNA sequencing techniques. *Quarterly Reviews of Biophysics* 35: 169–200.
- Gao, Q., G. Yue, W. Li, J. Wang, J. Xu, and Y. Yin. 2012. Recent progress using high-throughput sequencing technologies in plant molecular breeding. *Journal of Integrative Plant Biology* 54 (4): 215–227.
- Garcia-Mas, J., *et al.* [34 authors]. 2012. The genome of melon (*Cucumis melo* L.). *Proc. Nat. Acad. Sci. USA* 109: 11872–11877.
- Goff, S.A. *et al.* [55 authors]. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* 296: 92–100.
- Gupta, P.K., S. Rustgi, and R.R. Mir. 2008. Array-based high-throughput DNA markers for crop improvement. *Heredity* 101: 5–18.
- Hu, T.T., P. Pattyn, E.G. Bakker, J. Cao, J.F. Cheng, and R.M. Clark. 2011. *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. *Nat. Genet.* 43: 476–481.
- Huang, S. *et al.* [85 authors]. 2009. The genome of the cucumber, *Cucumis sativus* L. *Nat. Genet.* 41: 1275–1281.
- Huang, X. *et al.* [28 authors]. 2010. Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* 42: 961–967.
- Jallion, O., *et al.* [55 authors]. 2007. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* 449: 463–467.
- Kling, J. 2005. The search for sequencing through bred. *Nat. Biotechnol.* 23: 1333–1335.
- Long, S.P. and D.R. Ort. 2010. More than taking the heat: crops and global change. *Curr. Opin. Pl. Biol.* 13: 240–247.
- Metzker, M. 2010. Sequencing technologies. The next generation. *Nat. Rev. Genet.* 11: 31–46.
- Ming, R. *et al.* [92 authors]. 2008. The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nat. Lett.* 24: 991–997.
- Olsvik, O., J. Whalberg, B. Petterson, M. Uhlen, T. Popovic, I.K. Wachmuth, and P.I. Fields. 1993. Use of automated sequencing of polymerase chain reaction-generated amplicons to identify three types of cholera toxin subunit B in *Vibrio cholerae* O1 strains. *Journal of Clinical Microbiology* 31: 22–25.
- Paterson, A.H. *et al.* [45 authors]. 2009. The *Sorghum bicolor* genome and the diversification of grasses. *Nature* 457: 551–556.
- Paterson, A.H., M. Freeling, H. Tang, and X. Wang. 2010. Insights from the comparison of plant genome sequences. *Annu. Rev. Plant Biol.* 61: 349–372. doi: 10.1146/annurev-arplant-042809-112235.
- Rent, Y., Z. Hong, Q. Kou, J. Jiang, S. Guo, H. Zhang, W. Hou, X. Zou, H. Sun, G. Gong, A. Levi, and Y. Xu. 2012. A high resolution

- genetic map anchoring scaffolds of the sequenced watermelon genome. *PLoS One* 7: e29453.
- Sanger, F., S. Nicklen, and A.R. Coulson. 1997. DNA sequencing with chain-terminating inhibitors. *Proc. Nat. Acad. Sci. USA* 74: 5463–5467. doi:10.1073/pnas.74.12.5463
- Sato, S., *et al.* [29 authors]. 2008. Genome structure of the legume, *Lotus japonicus*. *DNA Res.* 15: 227–239.
- Satyawan, D., H. Rijzaani, and I.M. Tasma. 2014. Characterization of genomic variation in Indonesian soybean (*Glycine max*) varieties using next-generation sequencing. *Plant Genet. Resources* 12: S109-S113.
- Schmutz, J. *et al.* [44 authors]. 2010. Genome sequence of the paleopolyploid soybean. *Nature* 463: 178–183.
- Schnable, P.S. *et al.* [152 authors]. 2009. The B73 maize genome: complexity, diversity, and dynamics. *Science* 326: 1112–1115.
- Shendure, J., R.D. Mitra, C. Varma, and G.M. Church. 2004. Advanced sequencing technologies: methods and goals. *Nat. Rev. Genet.* 5: 335–344.
- Shulaev, V. *et al.* [72 authors]. 2011. The genome of woodland strawberry (*Fragaria vesca*). *Nat. Genet.* 43: 109–116
- Singh, R. *et al.* [87 authors]. 2013. Oil palm genome sequence reveals divergence of interfertile species in Old and New worlds. *Nature* 500: 335–339.
- Tasma, I.M., D. Satyawan, H. Rijzaani, D.W. Utami, P. Lestari, dan I. Rosdianti. 2012. Pembentukan empat peta genetik sawit, jarak pagar, padi, dan kedelai, serta identifikasi marka SNP kakao dan sapi. Laporan Akhir Penelitian APBN 2012. BB Biogen, Bogor. 67 hlm.
- Tasma, I.M., D. Satyawan, H. Rijzaani, P. Lestari, and Rubiyo. 2015a. SNP discovery based on whole genome sequences of five Indonesian cacao varieties. Manuscript in preparation.
- Tasma, I.M., H. Rijzaani, D. Satyawan, P. Lestari, D.W. Utami, I. Rosdianti, A.R. Purba, E. Mansyah, A. Sutanto, R. Kirana, Kusmana, A. Anggraeni, M. Pabendon, and Rubiyo. 2015b. Next-gen-based DNA marker development of several important crop and animal species. Manuscript Presented at the 13th SABRAO Congress and International Conference, 14–16 September 2015, Bogor, Indonesia. 8 pp.
- The Tomato Genome Consortium. 2012. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485: 635–641.
- Tuskan, G.A. *et al.* [111 authors]. 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313: 1596–1604.
- Upadhyaya, H.D., S. Sharma, K.N. Reddy, S.K. Singh, S. Singh, R.K. Varshney, and C.L.L. Gowda. 2013. Utilization of cultivated and wild germplasm in crop improvement. 3rd International Symposium on Genomics of Plant Genetic Resources, Jeju, Korea, 16–19 April, 2013. p. 18.
- Van, K., K. Rastogi, K.H. Kim, and S.H. Lee. 2013. Next-generation sequencing technology for crop improvement. *SABRAO J. Breed. Genet.* 45(1): 84–99.
- Varshney, R.K., S.N. Nayak, G.D. May, and J.A. Jackson. 2009. Next-generation sequencing technologies and their implications for crop genetics and breeding. *Trends in Biotechnol.* 9: 522–530.
- Varshney, R.K. *et al.* [31 authors]. 2012. Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Nat. Biotechnol.* 30: 83–89. doi:10.1038/nbt.2022 .
- Velasco, R. *et al.* [81 authors]. 2010. The genome of the domesticated apple (*Malus × domestica* Borkh.). *Nat. Genet.* 4: 833–839. doi:10.1038/ng.654
- Venter, J.C., S. Levy, T. Stockwell, K. Remington, and A. Halpern. 2003. Massive parallelism, randomness and genomic advances. *Nat. Genet.* 33: 219–227.
- Wang, X. *et al.* [91 authors]. 2011. The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* 43: 1035–1039.
- Wóycicki, R. *et al.* [76 authors]. 2011. The genome sequence of the North-European cucumber (*Cucumis sativus* L.) unravels evolutionary adaptation mechanisms in plants. *PLoS ONE* 6 (7): e22728. doi: 10.1371/journal.pone.0022728.
- Xu, X., *et al.* [94 authors]. 2011. Genome sequence and analysis of the tuber crop potato. *Nature* 475: 189–195.
- Young, N.D. *et al.* [56 authors]. 2011. The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature* 480: 520–524.
- Yu, J. *et al.* [85 authors]. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* 296: 79–92.