

MARKERMART@CMDV: Repositori maklumat genomik dan genetik bagi aplikasi pembangunan penanda molekul dan genomik berfungsi

(MARKERMART@CMDV: Repository of genomics and genetics information for application in the development of molecular markers and functional genomics)

Rabiatul Adawiah Zainal Abidin, Azrin Ahmad, Muhammad Fairuz Yusof, Norzihan Abdullah dan Umi Kalsom Abu Bakar

Pengenalan

Penggunaan teknologi maklumat dan komunikasi (ICT) bagi analisis dan pengurusan maklumat genomik dan genetik sama ada berbentuk jujukan, teks atau imej telah digunakan secara meluas. Antara kekangan yang menyebabkan penyelidik menggunakan ICT dalam menguruskan maklumat adalah seperti maklumat yang bertaburan, maklumat dengan format berbeza dan pencarian maklumat mengambil masa yang lama.

Pusat Penemuan dan Pengesahan Penanda Molekul (CMDV) aktif menjalankan penyelidikan yang menghasilkan data dan maklumat bertruput tinggi bagi aplikasi pembangunan penanda molekul dan genomik berfungsi. Misalnya, CMDV menjana data penanda molekul iaitu polimorfisme nukleotida tunggal (SNP) dan mikrosatelit (SSR) bagi tujuan kajian diversiti, cap jari DNA serta perkhidmatan kepada pelanggan CMDV. Penjujukan genom buah-buahan tropika seperti betik dan durian juga sedang dijalankan untuk pembangunan cap jari DNA manakala penjujukan transkriptomik buah duku dan tembikai telah dimulakan bagi merungkai gen-gen, tapak jalan dan kandungan bahan aktif yang terlibat dengan pelbagai ciri yang diingini seperti ciri kualiti dan rintang penyakit.

Penghasilan maklumat genomik dan genetik melalui kedua-dua aktiviti penyelidikan tersebut adalah bertruput tinggi dan maklumat adalah dalam pelbagai bentuk. Oleh itu, pembangunan sistem pangkalan data genomik dan genetik yang sistematik dan komprehensif sangat penting supaya maklumat tersebut boleh diarkib secara terurus dan diakses dengan mudah bagi pengguna seperti ahli biologi molekul, ahli genetik dan pembiak baka. Ini secara tidak langsung akan membantu pengguna meningkatkan produktiviti penyelidikan agrobioteknologi.

CMDV telah mereka bentuk dan membangunkan MARKERMART@CMDV sebagai pangkalan data utama yang menyimpan maklumat genomik dan genetik padi, buah-buahan (betik, pisang, nanas, mangga, durian dan duku) serta

sayuran terpilih (tomato, cili, ubi keledek dan jagung) bagi kegunaan pembangunan cap jari DNA, lokus ciri kuantitatif (QTL), kajian diversiti dan pemetaan peta fizikal.

Untuk membangunkan pangkalan data ini, satu penyelidikan mengenai pemilihan model pangkalan data dan keperluan fungsi pencarian yang diinginkan oleh pengguna dalam kajian penanda molekul dan genomik berfungsi telah dijalankan. Buat masa ini MARKERMART@CMDV mengumpul dan menggabungkan maklumat genomik dan genetik yang dihasilkan oleh makmal CMDV dan daripada pangkalan data umum.

Metodologi pembangunan MARKERMART@CMDV

Jenis maklumat genomik dan genetik yang terlibat ialah jujukan genom, nama penanda molekul SNP dan mikrosatelit, gen, tapak jalan, ciri fenotip, QTL, ciri serta janaplasma daripada padi, buah-buahan tropika dan sayur-sayuran terpilih. Setelah melalui beberapa siri pembersihan maklumat, barulah ia dimuat naik dalam MARKERMART@CMDV. Maklumat genomik dan genetik tersebut adalah berbentuk teks dan jujukan.

Kerangka MARKERMART@CMDV telah dibina berdasarkan enjin Wordpress (<http://wordpress.com>) yang memberi pautan kepada tiga jenis perisian bioinformatik iaitu BioMart (<http://biomart.org>), JBrowse (<http://jbrowse.org>) dan SequenceServer (<http://www.sequenceserver.com>) manakala MySQL telah dipilih sebagai sistem pengurusan pangkalan data (*Gambar 1*). Kesemua perisian tersebut merupakan sumber terbuka, mudah dibangunkan, mempunyai nilai-nilai sekuriti dan sangat aktif digunakan oleh komuniti penyelidik dan pentadbir sistem.

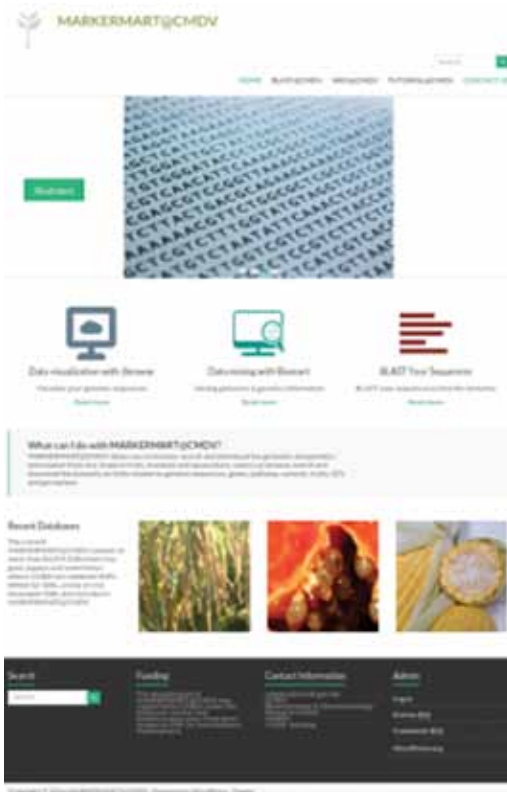
Buat masa ini, MARKERMART@CMDV beroperasi melalui pelayan komputer berasaskan sistem operasi Linux. Sehingga kini, MARKERMART@CMDV mengandungi 11 pangkalan data manakala saiz terkini pangkalan data MARKERMART@CMDV ialah 500 MB.

Deskripsi MARKERMART@CMDV

Pengurusan maklumat genomik dan genetik berbentuk teks adalah dengan menggunakan perisian BioMart yang merupakan sistem pengurusan maklumat yang dibangunkan dengan menggunakan skrip Java dan bersesuaian mengendalikan data-data biologi berkadar tinggi dalam bentuk teks. Sehingga kini, terdapat lebih kurang 40 pangkalan data biologi awam yang menggunakan aplikasi BioMart sebagai sistem pengurusan data (<http://www.biomart.org/notice.html>). Versi BioMart yang digunakan oleh MARKERMART@CMDV ialah BioMart 0.7. Antara maklumat genomik dan genetik yang boleh dicapai adalah seperti dalam *Jadual 1*.

Jadual 1. Antara maklumat genomik dan genetik yang boleh dicapai

Set data	Deskripsi
Padi	Panel 1536 SNP bagi 47 varieti padi QTL dengan 4 ciri (kandungan amilosa, wangi, kepanjangan dan rintang penyakit RTBV)
Betik	800 putatif SNP daripada varieti Eksotika dan Sekaki Gen rintang penyakit
Pisang	10,135 putatif mikrosatelit daripada genom pisang Musa acuminata cv Pahang 96 primer mikrosatelit
Mangga	300 putatif mikrosatelit daripada transkriptomik Mangifera indica cv Langra
Nanas	54,964 putatif mikrosatelit daripada genom nanas kultivar F153 176 putatif mikrosatelit daripada genom nanas kultivar MD2 8 primer mikrosatelit
Duku	300 putatif mikrosatelit daripada gen-gen terpilih Senarai gen dan tapak jalan
Tembikai	Panel 800 SNP bagi 24 aksesii tembikai Peta fizikal, QTL dengan 10 ciri Senarai gen dan tapak jalan
Tomato	20,209 putatif mikrosatelit daripada genom tomato
Cili	36,721 putatif mikrosatelit daripada genom cili
Ubi keledek	96 primer mikrosatelit daripada penanda jujukan terekspres (EST) ubi keledek
Jagung	325 putatif mikrosatelit daripada genom jagung



Gambar 1. Antara muka grafik utama MARKERMART@CMDV

Pengguna boleh melakukan pencarian serta memuat turun maklumat genomik dan genetik yang dikehendaki dengan begitu mudah (*Gambar 2*). Data dan maklumat boleh dimuat turun mengikut format seperti .xls, .csv dan .txt.

Antara penggunaan lain MARKERMART@CMDV ialah visualisasi data berbentuk jujukan digambarkan dengan menggunakan perisian JBrowse. Melalui JBrowse, pengguna dapat melihat lokasi SNP dan mikrosatelit pada genom dengan lebih terperinci (*Gambar 3*).

Selain itu, pengguna turut berpeluang menjalankan analisis pencarian persamaan dengan menggunakan perisian BLAST yang telah dipautkan pada MARKERMART@CMDV (*Gambar 4*). Pengguna perlu memasukkan jujukan DNA atau protein dan memilih pangkalan data padi, buah-buahan atau sayur-sayuran yang diinginkan oleh mereka. Pencarian persamaan ini membolehkan pengguna mengetahui maklumat gen atau protein yang boleh diramal dalam jujukan suatu organisma.



Gambar 2. Pengguna boleh memuat turun maklumat mengikut format .xls, .csv dan .txt



Gambar 3. Visualisasi posisi jujukan padi, buah-buahan dan sayur-sayuran pada genom masing-masing dengan menggunakan perisian JBrowse



Gambar 4. Pencarian persamaan jujukan DNA dan protein menggunakan perisian BLAST turut dimuatkan dalam MARKERMART@CMDV

Rasional pembangunan MARKERMART@CMDV

MARKERMART@CMDV dibangunkan untuk memudahkan pengguna menguruskan maklumat genomik dan genetik yang diperlukan bagi aplikasi pembangunan penanda molekul dan genomik berfungsi. Melalui penggunaan MARKERMART@CMDV, maklumat genomik dan genetik yang telah dihasilkan oleh penyelidik CMDV dapat disepadukan dalam satu platform yang memudahkan pencarian dan pengumpulan data-data tersebut yang secara tidak langsung meningkatkan kecekapan pengurusan penyelidikan penanda molekul di CMDV.

Perancangan jangka masa panjang MARKERMART@CMDV adalah seperti menambah bilangan maklumat jujukan genomik dan genetik serta meningkatkan lagi fungsi MARKERMART@CMDV dengan mengintegrasikan beberapa perisian bioinformatik yang lain ke dalam sistemnya. MARKERMART@CMDV mampu berperanan sebagai repositori maklumat genomik dan genetik padi, buah-buahan tropika dan sayur-sayuran yang bertaraf nasional. Ini adalah berdasarkan kedudukan MARDI sebagai penjana utama janaplasm padi, buah-buahan tropika dan sayur-sayuran di Malaysia. Pembangunan MARKERMART@CMDV berteraskan sistem yang fleksibel dan kos yang rendah kerana menggunakan perisian-perisian sumber terbuka yang telah diperakui mempunyai nilai-nilai saintifik berdasarkan penerbitan pembangun perisian-perisian yang terlibat.

Penutup

Secara umumnya, MARKERMART@CMDV menyediakan maklumat berbentuk saintifik bagi membantu meningkatkan produktiviti penyelidikan penanda molekul dan pembaikbakaan tanaman. Pembangunan dan pengurusan MARKERMART@CMDV akan terus dikemas kini dan dikembangkan dari masa ke masa bagi menjamin kualiti dan nilai-nilai saintifik maklumat yang didepositkan.

Penghargaan

Kajian ini telah dijalankan di Pusat Penyelidikan Bioteknologi dan Nanoteknologi, MARDI dengan peruntukan dari pihak NKEA-EPP14: Pembangunan Industri Benih.

Bibliografi

- Damian, S., Syed, H., Benoit, B., Richard, H., Darin, L., Gudmundur, T. dan Arek, K. (2009). BioMart – biological queries made easy. *BMC Genomics* 10: 22
- Shirasawa, K., Isobe, S., Tabata, S. dan Hirakawa, H. (2014). Kazusa Marker Database: a database for genomics, genetics and molecular breeding in plants. *Breeding Science* 64: 264 – 271

Ringkasan

Artikel ini membincangkan MARKERMART@CMDV iaitu sistem pangkalan data yang dibangunkan bagi mengurus serta menyepadukan maklumat genomik dan genetik yang telah dihasilkan oleh CMDV dan daripada pelbagai pangkalan data umum. MARKERMART@CMDV berperanan sebagai platform perlombongan data kepada pengguna bagi aplikasi pembangunan penanda molekul dan genomik berfungsi.

Summary

This article discusses the MARKERMART@CMDV, a database management system developed by CMDV that used to manage and integrate genomics and genetics information produced by CMDV and also from public databases. MARKERMART@CMDV provides data mining platform for molecular biologist, geneticist and breeder for application in the development of molecular markers and functional genomics.

Pengarang

Rabiatul Adawiah Zainal Abidin
Pusat Penyelidikan Bioteknologi dan Nanoteknologi,
Ibu Pejabat MARDI, Serdang,
Persiaran MARDI-UPM, 43400 Serdang, Selangor
E-mel: rabiatul@mardi.gov.my

Azrin Ahmad, Muhammad Fairuz Yusof dan Norzihan Abdullah
Pusat Penyelidikan Bioteknologi dan Nanoteknologi,
Ibu Pejabat MARDI, Persiaran MARDI-UPM,
43400 Serdang, Selangor

Umi Kalsom Abu Bakar
Pejabat Ketua Pengarah MARDI, Ibu Pejabat MARDI, Serdang,
Persiaran MARDI-UPM, 43400 Serdang, Selangor