

BAB I

PENDAHULUAN

1.1. Latar Belakang

Menurut data BPS (2015) dalam Kemendag(2017), dengan luas lahan sebesar 3,8 juta hektar, pada tahun 2014 produksi jagung dalam bentuk pipilan kering mencapai 19,03 juta ton atau mengalami kenaikan sebesar 2,81% dibandingkan pada tahun 2013 (18,51 juta ton). Kenaikan produksi terjadi baik di Pulau Jawa maupun diluar Pulau Jawa pada periode sekitar Mei-Agustus dan September-Desember 2014 karena adanya kenaikan produktivitas jagung sekitar 2,87% per tahun. Melalui berbagai langkah yang telah dilakukan dan akan ditempuh, Kementrian Pertanian menargetkan produksi jagung tahun 2015 mencapai 20,3 juta ton.

Swastika et al. (2004) dan Azria (2006), dalam Yulita & Naiola (2013), menulis bahwa jagung merupakan sumber pangan biji-bijian kedua terpenting setelah beras, pemerintah Indonesia bahkan telah mencanangkan swasembada jagung pada tahun 2007. Jagung telah menjadi sumber karbohidrat utama di beberapa provinsi di Indonesia, yaitu Jawa Timur, Nusa Tenggara Timur, Sulawesi Utara, Sulawesi Selatan dan Lampung (Rachman , 2003). Produksi jagung di Indonesia mengalami kenaikan rata-rata 4,07% pertahun sejak tahun 1970 hingga mencapai 1,4 juta ton pada tahun 2006. Kenaikan produksi ini terutama disebabkan oleh berkembangnya teknologi dalam merakit varietas unggul. Di Indonesia hingga saat ini telah dilepas 33 varietas unggul dan menyebar cukup luas di Indonesia. Dengan banyaknya penggunaan varietas unggul oleh petani maka ras lokal bisa terdesak atau bahkan bisa musnah (Yulita & Naiola, 2013).

Upaya perbaikan sifat tanaman varietas jagung dengan meningkatkan keragaman genetiknya perlu dilakukan. Seperti telah diketahui, modal dasar pemuliaan tanaman adalah adanya keragaman yang luas. Sobir (2007) dalam

Sain (2016) menyatakan bahwa dengan adanya variabilitas yang luas proses seleksi dapat dilakukan secara efektif karena akan memberikan peluang yang lebih besar untuk diperoleh karakter-karakter yang diinginkan.

Keragaman varietas jagung merupakan pondasi dasar dalam program pemuliaan tanaman untuk menghasilkan varietas unggul. Pada prinsipnya varietas unggul dirakit dengan melalui tahapan evaluasi pemilihan plasma nutfah varietas jagung pembentuk galur-galur murni, pengujian multi lingkungan, pelepasan varietas baru dan perbanyakan benih (Febriani, et al., 2008). Dengan berkembangnya ilmu pengetahuan dan teknologi yang berbasis pada DNA, diperlukan suatu teknologi untuk menangani itu. Salah satu teknologi itu yaitu bioinformatika.

Bioinformatika adalah gabungan antara ilmu biologi dan ilmu teknik informasi (TI). Pada umumnya bioinformatika didefinisikan sebagai aplikasi dari alat komputasi dan analisa untuk menangkap dan menginterpretasikan data-data biologi. Ilmu ini merupakan ilmu baru yang yang merangkul berbagai disiplin ilmu termasuk ilmu komputer, matematika dan fisika, biologi, dan ilmu kedokteran, dimana kesemuanya saling menunjang dan saling bermanfaat satu sama lainnya. Istilah bioinformatik mulai dikemukakan pada pertengahan era 1980-an untuk mengacu pada penerapan komputer dalam biologi. Namun demikian, penerapan bidang-bidang dalam bioinformatika (seperti pembuatan basis data dan pengembangan algoritma untuk analisis sekuens biologis) sudah dilakukan sejak tahun 1960-an (Fatchiyah, 2009). Salah satu teknologi yang berkembang dalam bioinformatik adalah teknologi *microarray*.

Microarray merupakan teknologi dalam bidang Biologi Molekuler dan Medis yang dapat digunakan untuk melihat perbedaan ekspresi gen. Selain itu, *microarray* dapat digunakan untuk mendeteksi *single nucleotide polymorphism* (SNP) and *genotyping*. Teknologi ini memanfaatkan kumpulan array yang berjumlah ribuan yang berisi nukleotida DNA yang berfungsi sebagai probe. Hibridisasi antara probe dan target (cDNA atau cRNA) dideteksi dengan menggunakan target yang dilabel fluoresen.

Karena array yang digunakan terdiri atas ribuan probe, eksperimen *microarray* dapat dikatakan sebagai tes genetik yang dilakukan secara paralel. Informasi yang dihasilkan sangat detail dan menyeluruh pada genom pada tingkat transkripsi gen. Sehingga, proses biologi yang melibatkan regulasi gen bisa dianalisis dengan lebih baik. Informasi yang diperoleh dari hasil *microarray* telah dimanfaatkan untuk berbagai aplikasi spesifik seperti diagnosis penyakit, penemuan obat-obatan, pengelompokan ekspresi gen yang terlibat dalam organogenesis, cekaman dan interaksi dengan mikro organisme (Pertanian, 2010).

Saat ini teknologi sekuensing DNA telah berkembang dengan munculnya *Next Generation Sequencing* (NGS). *Next Generation Sequencing* (NGS) adalah terminologi yang digunakan proses *sequencing* modern. Ini dideskripsikan sebagai sejumlah perbedaan teknologi *sequencing* modern dengan revolusi penelitian genomik dan biologi molekuler. Teknik ini adalah *illumina sequencing*, *roche 454 sequencing*, *ion protons sequencing* dan *sequencing oligo litation detektion*. Sistem NGS sangat cepat dan lebih murah. Empat metode *sequencing* DNA utama digunakan, dinamai *pyrosequencing*, *sequencing* dengan ligasi, *sequencing* dengan sintesis dan ion *semiconductor sequencing*. Sejumlah besar rantai DNA dan RNA dapat di *sequencing* paralel. Ini membuat *sequencing* genom organisme dalam waktu singkat, tidak seperti *sequencing* Sanger yang memakan waktu lebih lama (Santiko, 2011).

Sekuensing genom dengan menggunakan teknologi NGS telah diterapkan pada beberapa tanaman perkebunan dan pertanian seperti kakao, kelapa sawit, kopi, karet, tebu, padi dan jagung. Penggunaan teknologi NGS tidak hanya terbatas pada sekuensing genom saja, namun juga telah menjujupada tahap aplikasi seperti pengembangan penanda molekuler, deteksi ekspresi gen skala besar menggunakan transkriptom, studi ekologi melalui varietas tanaman perkebunan, identifikasi karakter morfologi penting, dan pengelompokkan organ tumbuhan dengan karakteristik yang sama.

Penelitian yang dilakukan oleh Fiannaca et al., (2015) telah dilakukan menggunakan RNA-seq dengan teknologi NGS (*Next Generation Sequence*) pada data kanker payudara. Teknologi NGS digunakan untuk menganalisis perbedaan ekspresi gen kanker. Dengan menggunakan algoritma *machine learning* membantu dalam menemukan pola RNA-seq kanker payudara pada tingkat molekuler. Identifikasi pola gen dengan karakteristik yang sama dilakukan dengan analisis *biclustering* (Pontez et al., 2015).

Dalam upaya mendukung program varietas unggul, teknologi baru seperti bioinformatika sangat dibutuhkan untuk memaksimalkan keberhasilan perakitan suatu varietas tanaman. Contoh bioinformatika yang dihasilkan sampai saat ini adalah tanaman varietas unggul yang diperoleh melalui proses pembuatan varietas unggul dengan menggunakan teknologi *biclustering*.

Pontez et al., (2015) menyatakan bahwa *probabilistics model* adalah pembuatan model secara statistik yang menggambarkan data berdasarkan teori peluang. Salah satu algoritma *biclustering* dengan tipe *probabilistics model* adalah model *plaid*.

Penelitian yang dilakukan oleh Lazzeroni & Owen (2002) menyatakan bahwa model *plaid* adalah suatu metode untuk mengeksplorasi data dengan analisis multivariat. Pendekatan ini, diterapkan pada matriks gen ekspresi yang kemudian di sesuaikan *bicluster*.

Beberapa hasil penelitian telah membuktikan analisis model *plaid* sukses untuk mengidentifikasi penyakit. Ardaneswari et al. (2017) menyebutkan bahwa memanfaatkan analisis model *plaid* untuk mengidentifikasi pengelompokan gen tumor *carcinoma* dan gen tumor *adenoma*. Pada penelitian tersebut ditemukan tiga *bicluster* pada gen tumor *carcinoma* dan empat *bicluster* terbentuk dari gen tumor *adenoma*. Berdasarkan contoh literatur terdahulu tersebut membuktikan dengan menggunakan analisis model *plaid* sangat potensial untuk mendeteksi pengelompokan data pada data *maize growth stage*.

Sebuah pengembangan analisis model *plaid* dilakukan oleh Kasim et al. (2017) dengan menyediakan *package biclust* pada bahasa pemrograman R. *Package* tersebut membantu dalam pembentukan visualisasi *bicluster* dan analisis model *plaid*. Selain *open source*, bahasa pemrograman R sangat potensial digunakan di masa yang akan datang, karena memungkinkan para *user* untuk saling melengkapi *function* yang dibutuhkan pada analisis model *plaid*.

Maka dari itu, sesuai dengan hal-hal yang telah disampaikan sebelumnya, penelitian kali ini ingin mendeteksi *bicluster* pada data *maize growth stage* sebagai bahan kajian peningkatan kualitas tanaman jagung menggunakan analisis model *plaid* dengan alat bantu bahasa pemrograman R.

1.2. Rumusan Masalah

Berdasarkan dengan latar belakang yang telah dipaparkan, maka penulis dapat merumuskan beberapa masalah yang hendak diselesaikan pada penelitian ini, antara lain:

1. Bagaimana hasil eksplorasi data *Maize Growth Stage* menggunakan bahasa pemrograman R?
2. Bagaimana mengidentifikasi pola data *Maize Growth Stage* menggunakan *bicustering* dengan algoritma *Plaid* menggunakan bahasa pemrograman R?
3. Bagaimana mengimplementasikan hasil kajian algoritma *Plaid* dengan bahasa pemrograman R tersebut pada program pemuliaan tanaman untuk meningkatkan di Indonesia?

1.3. Batasan Masalah

Data yang digunakan pada penelitian ini adalah data *maize growth stage*. Pada penelitian ini menggunakan bahasa pemrograman R versi 3.5.0 dengan *package R/RcmdrPlugin.BiclustGUI* untuk analisis *bicustering*.

1.4. Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Mengetahui informasi hasil eksplorasi data *Maize Growth Stage* dengan bahasa pemrograman R.
2. Mengetahui hasil identifikasi pola data *Maize Growth Stage* menggunakan *bicustering* dengan metode *Plaid* menggunakan bahasa pemrograman R.
3. Mengetahui implementasi hasil kajian *bicustering* dengan bahasa pemrograman R tersebut pada program pemuliaan tanaman untuk meningkatkan kualitas jagung di Indonesia.

1.5. Manfaat Penelitian

1. Mengetahui hasil pengaplikasian *bicustering* algoritma *plaid* pada data *maize growth stage*.
2. Dapat mengetahui hasil pengelompokkan terhadap data uji dengan menggunakan metode *bicustering* algoritma *plaid*.