

# BAB I PENDAHULUAN

## 1.1 Latar Belakang

Salah satu masalah besar yang menjadi perhatian dalam informasi molekular dalam aspek taksonomi biologi adalah bagaimana cara terbaik dalam pembacaan *DNA barcoding*. Terdapat dua tujuan utama dimana *DNA barcode* diterapkan, tujuan pertama adalah untuk membedakan antar spesies makhluk hidup dan tujuan kedua adalah penggunaan data *DNA* untuk menemukan spesies baru (DeSalle et al. 2005). *DNA barcoding* merupakan teknik dalam mengidentifikasi spesies dengan menggunakan sekuens *DNA* dari sebagian kecil genom, dengan tujuan berkontribusi pada berbagai studi ekologi dan konservasi (Lahaye et al. 2008). Diagnosis pada skala molekuler melalui *DNA barcoding* memberikan alternatif identifikasi tanaman yang cepat, akurat, dan tidak ambigu dalam proses identifikasinya (Virgilio et al. 2012). Identifikasi tanaman dengan *DNA barcoding* memerlukan jumlah sampel yang sedikit dan pada prinsipnya *DNA* dapat diambil dari bagian tanaman apa saja.

Dalam *DNA barcoding* terdapat beberapa proses di dalamnya yang masih menggunakan cara konvensional terutama pada tahap *similarity check* yaitu tahap untuk melihat persamaan antar sekuens *DNA*. Pada saat ini khususnya dalam tahap *similarity check* belum ada program atau sejenisnya yang menggunakan komputer yang dapat mengakomodir tahap tersebut. Sehingga pada prosesnya masih menggunakan cara yang konvensional yaitu dengan cara melihat sekuens *DNA* lalu menandai daerah-daerah mana yang merupakan daerah berpotensi yang akan dijadikan sebagai *barcode*. Oleh sebab itu tingkat akurasi masih rendah dan masih memerlukan waktu yang cukup lama (Hidayat, *unpublished manuscript*).

Pendekatan yang mengarah kepada deteksi motif *DNA* menghadapi dua jenis tantangan. Pertama, Faktor transkripsi *DNA* dimana terikatnya protein pada deret dan motif struktur *DNA* tertentu yang biasanya terletak pada hulu gen target, biasanya memiliki struktur yang pendek. Kedua, motif *DNA* lemah untuk dikonservasi karena adanya evolusi dan mutasi. Melihat daripada masalah ini, baik metode perbandingan string sederhana maupun *exhaustive search* semua kombinasi

dari beberapa metode tidak dapat secara efektif memberikan identifikasi faktor transkripsi yang akurat (Yang & Wang, 2008).

Salah satu penelitian yang dilakukan oleh Yang dan Chuang (2011) tentang *motif discovery* DNA, dan dapat disimpulkan bahwa penggunaan kemampuan *Ant Colony Optimization* (Dorigo & Caro, 1999) sebagai *stochastic search* dan prosedur maksimalisasi EM menunjukkan penggunaan keduanya dapat dikatakan sangat menjanjikan dalam masalah penemuan motif DNA, atau DNA *motif discovery*. Dengan menghasilkan *F-score* atau nilai rata-rata hasil melebihi *tools* GAME dan GALF. GAME merupakan perangkat lunak atau *tools* yang digunakan dalam *motif discovery* pada sekuens DNA menggunakan algoritma genetika (Wei & Jensen, 2006). GALF merupakan singkatan dari *Genetic Algorithm Local Filtering*, serupa dengan GAME, GALF merupakan *tools* yang digunakan untuk *motif discovery* pada sekuens DNA menggunakan algoritma genetik melalui lokal *filtering*. Namun pada penelitian ini masih memiliki kelemahan, yaitu masih dibutuhkannya *fitness function* yang lebih efektif untuk membedakan situs pengikatan yang palsu dengan yang aslinya, dan hanya mampu melakukan satu *binding site* per sekuennya.

Penelitian lain Hardin dan Rouchka (2005) mengenai *motif discovery* DNA, memiliki perbedaan dengan Yang dan Chuang (2011) tentang *motif discovery*, yaitu pada penelitian ini menggunakan *particle swarm optimization* (Kennedy & Eberhart, 1995) dan sama-sama menggunakan *Expectation Maximization* (Dempster et al, 1977). Pada penelitian ini disebutkan bahwa meskipun menggunakan pemanfaatan topologi lingkungan, *particle swarm optimization/Expectation Maximization* masih berhenti pada optimum lokal. Penelitian ini bermaksud menyelidiki sifat dari gradien (jika ada) yang dapat digunakan untuk menggerakkan partikel ke arah yang lebih optimal.

Penelitian yang akan dilakukan oleh penulis akan menggunakan sebuah algoritma optimasi yaitu Algoritma *Spiral Dynamic*. Algoritma *Spiral Dynamic* dikembangkan oleh Tamura dan Yasuda (2011) dari Universitas Tokyo Metropolitan, Jepang. Algoritma *Spiral Dynamic* menggunakan pendekatan metaheuristik yang berdasarkan kepada analogi dari fenomena spiral di alam yang berkonsep sederhana dan mudah dalam implementasinya.

Algoritma Optimisasi Spiral Dynamic merupakan sebuah pendekatan metaheuristik dua dimensi yang baru dan terinspirasi dari fenomena spiral yang ada pada alam. Fitur umum dari algoritma ini ialah melakukan strategi pencarian yang efektif dalam pendekatan metaheuristik dan juga dapat membangun algoritma optimisasi yang baru berdasarkan model diskret. Keluaran dari simulasi untuk masalah benchmark dua dimensi ini menghasilkan keefektivitasan dan cukup potensial dari metode yang diusulkan (Tamura & Yasuda, 2011).

Pada saat ini penggunaan algoritma optimisasi, salah satunya adalah algoritma *Spiral Dynamics* (Tamura & Yasuda, 2011) khususnya pada pencarian motif *DNA barcode* pada DNA masih sedikit sekali digunakan ataupun bahkan belum ada yang menggunakan, maka untuk penggunaan algoritma *Spiral Dynamic*, khususnya pada *motif discovery DNA barcode* atau lebih khususnya pada tahap *similarity check* di dalam *DNA barcoding* belum ditemukan sama sekali. Algoritma *Spiral Dynamics* juga merupakan sebuah algoritma yang baru muncul pada tahun 2011. Pada kasus *DNA Barcoding* dalam tahap *similarity check* untuk prosesnya digunakan metode *hamming distance* yaitu mencari perbedaan di dalam sebuah *string* atau karakter yang dibandingkan. Algoritma *spiral dynamic* ini berperan untuk mengoptimasi hasil dari *hamming distance* tersebut ke angka minimum, yang berarti kecocokan antar karakter mendekati sama.

Maka dari itu penulis tertarik untuk membuat penelitian ini, dan diharapkan dengan adanya penelitian ini dapat memberikan alternatif terkait masalah *DNA barcoding* dan juga yang diharapkan dari penelitian ini memberikan hasil dengan akurasi yang tinggi pada motif *barcode* yang ditemukan dan kehandalan yang lebih tinggi dibandingkan dengan metode yang sudah ada sebelumnya.

## 1.2 Rumusan Masalah

Rumusan masalah dari latar belakang yang telah dijelaskan pada sub bab sebelumnya adalah sebagai berikut.

1. Bagaimana model untuk algoritma optimasi *spiral dynamic* pada sekuens *DNA* dalam kasus *DNA barcoding*?
2. Bagaimana penerapan algoritma optimasi *Spiral Dynamic* dapat bekerja pada R?

3. Bagaimana hasil keluaran dari program yang telah didesain dan dilakukan implementasi?

### 1.3 Tujuan

Berdasarkan permasalahan sebelumnya, maka tujuan dilakukannya analisis ini adalah sebagai berikut.

1. Merancang model untuk algoritma optimasi *spiral dynamic* pada sekuens DNA dalam kasus DNA *barcoding*.
2. Melakukan implementasi model algoritma optimasi *spiral dynamic* pada sekuens DNA dalam kasus DNA *barcoding* dalam bahasa pemrograman R.
3. Melakukan eksperimen setelah program selesai dibuat, lalu menganalisis hasil keluaran setelah dilakukan optimisasi dengan mengacu kepada aspek akurasi dan kecepatan.

### 1.4 Manfaat Penelitian

Terdapat beberapa manfaat dari hasil penelitian ini, yaitu.

1. Memberikan alternatif lain untuk DNA *barcoding* pada sekuens DNA menggunakan algoritma optimasi *spiral dynamic*.
2. Memudahkan para peneliti di bidang biologi untuk pendeteksian atau pencarian barcode DNA pada sekuens DNA.

### 1.5 Batasan Masalah

Batasan masalah yang terdapat pada penelitian ini yang bertujuan untuk lebih memfokuskan bahasan pada penelitian ini meliputi beberapa hal berikut ini.

1. Sekuens DNA yang dikenali yaitu nukleobasa seperti *adenine*(A), *guanine*(G), *cytosine*(C), *thymine*(T).
2. *Spiral Dynamic* yang dapat digunakan dimulai dari 2 dimensi dan seterusnya.

### 1.6 Sistematika

Berikut ini merupakan sistematika penulisan dalam proposal skripsi ini.

#### 1. BAB I: PENDAHULUAN

Dalam bab ini dijelaskan mengenai latar belakang dilakukannya penelitian mengenai penggunaan metode optimasi *spiral dynamics* untuk DNA *barcoding*,

rumusan masalah yang ada, tujuan yang ingin dicapai dari penelitian ini, batasan masalah, dan sistematika penulisan dalam penelitian ini.

## 2. BAB II: KAJIAN PUSTAKA

Bab ini menjelaskan atau menjabarkan tentang kajian pustaka atau landasan teori dari sumber-sumber ilmiah seperti buku, jurnal dan lainnya, yang mendukung seperti konsep dan pembahasan mengenai DNA, Bioinformatika, DNA *Barcoding*, *motif discovery*, algoritma *Spiral Dynamic* dan mengenai Bahasa pemrograman R.

## 3. BAB III: METODOLOGI PENELITIAN

Bab ini menjelaskan tentang segala hal seperti metode penelitian yang digunakan, objek penelitian, variabel, serta analisis dan kebutuhan sistem yang akan digunakan dalam penelitian ini.

## 4. BAB IV: TEMUAN DAN PEMBAHASAN

Pada bab ini dijelaskan hasil dari penelitian dan eksperimen yang telah dilakukan. Pertanyaan atau masalah yang dikemukakan dalam tema penelitian dibahas pada bab ini. Beberapa diantaranya seperti proses pengumpulan data sekuens DNA, implementasi sistem, desain eksperimen, dan hasil analisa.

## 5. BAB V: KESIMPULAN DAN SARAN

Pada bab ini berisi kesimpulan secara keseluruhan mengenai hasil eksperimen yang didapat dalam melakukan penerapan algoritma optimasi *spiral dynamic* untuk DNA *barcoding* menggunakan bahasa pemrograman R. Terdapat juga saran bagi peneliti untuk kedepannya atau pada penelitian selanjutnya.