

SISTEMATIKA DAN FILOGENETIKA MOLEKULER

Topik Hidayat dan Adi Pancoro

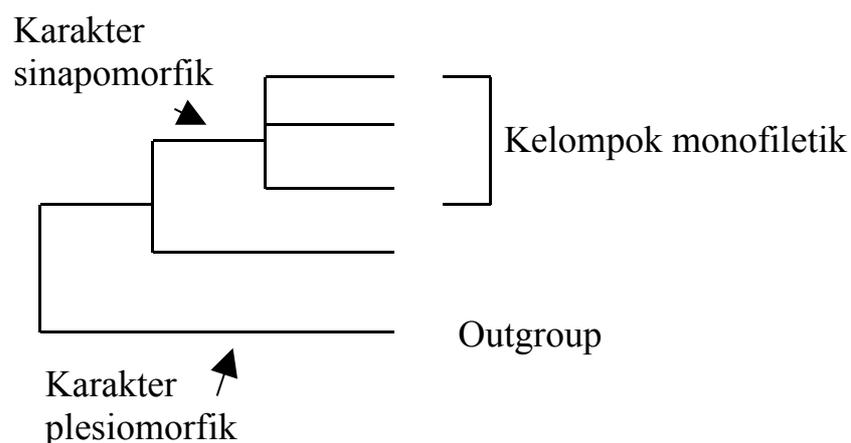
1. Apa yang ingin dicapai di dalam Sistematika?

Sistematika memiliki peran sentral di dalam Biologi dalam menyediakan sebuah perangkat pengetahuan untuk mengkarakterisasi organisme dan sekaligus merekognisinya dalam rangka memahami keanekaragaman. Secara fundamental, sistematika bertujuan untuk memahami dan mendeskripsikan keanekaragaman suatu organisme dan merekonstruksi hubungan kekerabatannya terhadap organisme lainnya, dan juga mendokumentasikan perubahan-perubahan yang terjadi selama evolusinya dan merubahnya ke dalam sebuah sistem klasifikasi yang mencerminkan evolusinya tersebut. Oleh karena itu, salah satu tugas yang penting dari sistematika adalah merekonstruksi hubungan evolusi (*evolutionary relationship*) dari kelompok-kelompok organisme biologi. Sebuah hubungan evolusi yang direkonstruksi dengan baik dapat digunakan sebagai landasan untuk melakukan penelitian-penelitian komparatif (*comparative investigations*) misalnya dalam bidang ekologi dan biogeografi.

Sampai saat ini ada dua pendekatan untuk merekonstruksi hubungan evolusi dari sebuah kelompok organisme biologi, yaitu fenetik dan kladistik. Kalau pendekatan pertama menaksir hubungan evolusi berdasarkan kepemilikan karakter atau ciri yang sama (*overall similarity*) dari anggota-anggota suatu kelompok, maka pendekatan kedua mendasari sebuah hubungan pada perjalanan evolusi karakter atau ciri dari setiap anggota suatu kelompok yang sedang dipelajari. Kladistik sering disebut atau ditulis di dalam literature ilmiah sebagai filogenetika dan merupakan pendekatan yang umum digunakan di dalam banyak penelitian sistematika.

Di dalam pendekatan filogenetika, sebuah kelompok organisme dimana

anggota-anggotanya memiliki banyak kesamaan karakter atau ciri dianggap memiliki hubungan yang sangat dekat dan diperkirakan diturunkan dari satu nenek moyang. Nenek moyang dan semua turunannya akan membentuk sebuah kelompok monofiletik. Dalam analisis filogenetika kelompok *outgroup* sangat dibutuhkan dan menyebabkan polarisasi karakter atau ciri, yaitu karakter apomorfik dan plesiomorfik. Karakter apomorfik adalah karakter yang berubah dan diturunkan dan terdapat pada *ingroup*, sedangkan karakter plesiomorfik merupakan karakter *primitive* yang terdapat pada *outgroup*. Karakter sinapomorfik adalah karakter yang diturunkan dan terdapat pada kelompok monofiletik. Gambar 1 memperlihatkan sebuah pohon kekerabatan sebagai hasil dari analisis filogenetika.



Gambar 1. Pohon kekerabatan dan polarisasi karakter dalam analisis filogenetika

Lebih lanjut, pohon filogenetika yang dihasilkan dapat diterjemahkan ke dalam sebuah sistem klasifikasi (sering disebut klasifikasi filogenetika), dan untuk itu hanya kelompok monofiletik yang dapat digunakan. Berbeda dengan klasifikasi tradisional (klasifikasi Linnaeus), klasifikasi filogenetika mengelompokkan satu

takson berdasarkan pengetahuan tentang hubungan filogenetika takson tersebut dengan takson lainnya, sehingga sistem klasifikasi yang dihasilkan bersifat obyektif.

2. Filogenetika Molekuler

Karakter morfologi telah lama digunakan dalam banyak penelitian filogenetika. Dengan pesatnya perkembangan teknik-teknik di dalam biologi molekuler, seperti PCR (*polymerase chain reaction*) dan sekuensing DNA, penggunaan sekuens DNA dalam penelitian filogenetika telah meningkat pesat dan telah dilakukan pada semua tingkatan taksonomi, misalnya famili, marga, dan species. Filogenetika molekuler mengkombinasikan teknik biologi molekuler dengan statistik untuk merekonstruksi hubungan filogenetika.

Pemikiran dasar penggunaan sekuens DNA dalam studi filogenetika adalah bahwa terjadi perubahan basa nukleotida menurut waktu, sehingga akan dapat diperkirakan kecepatan evolusi yang terjadi dan akan dapat direkonstruksi hubungan evolusi antara satu kelompok organisme dengan yang lainnya. Beberapa alasan mengapa digunakan sekuens DNA: (1) DNA merupakan unit dasar informasi yang mengkode organisme; (2) relatif lebih mudah untuk mengekstrak dan menggabungkan informasi mengenai proses evolusi suatu kelompok organisme, sehingga mudah untuk dianalisis; (3) peristiwa evolusi secara komparatif mudah untuk dibuat model; dan (4) menghasilkan informasi yang banyak dan beragam, dengan demikian akan ada banyak bukti tentang kebenaran suatu hubungan filogenetika.

Sekuens DNA telah menarik perhatian para praktisi taksonomi dunia untuk dijadikan karakter dalam penelitian filogenetika karena beberapa fakta. *Pertama*, sekuens DNA menawarkan data yang akurat melalui pengujian homologi yang lebih baik terhadap karakter-karakter yang ada. *Kedua*, sekuens DNA menyediakan banyak

character states karena perbedaan laju perubahan basa-basa nukleotida di dalam lokus yang berbeda adalah besar. Dan *ketiga*, sikuen DNA telah terbukti menghasilkan sebuah hubungan kekerabatan yang lebih alami (*natural*).

Sumber karakter DNA dapat diperoleh dari inti (nDNA), kloroplas (cpDNA), dan mitokondria (mtDNA). Tabel 1 memuat beberapa sistem gen dan genom yang telah digunakan dalam penelitian filogenetika molekuler.

Tabel 1 Beberapa sistem gen dan genom dalam studi filogenetika molekuler

Gen atau genom	Metode
DNA kloroplas	Analisis restriksi
Gen <i>rbcL</i> dari DNA kloroplas	Analisis sikuen DNA
Gen <i>matK</i> dari DNA kloroplas	Analisis sikuen DNA
DNA mitokondria	Analisis sikuen DNA
Gen RNA inti	Analisis sikuen DNA
Daerah ITS dari nrDNA	Analisis sikuen DNA
Kelompok gen repetitive:	
- Knob heterokromatin	Analisis sikuen DNA
- Gen CAB	Analisis sikuen DNA
- Gen <i>rbcS</i>	Analisis sikuen DNA
Gen kopi tunggal	Analisis sikuen DNA

Dari tabel di atas, para peneliti dihadapkan pada dua dasar pemilihan. *Pertama*, gen atau genom yang mana yang harus dipilih sebagai jawaban terbaik terhadap suatu pertanyaan filogenetika. Misalnya, masalah filogenetika pada tumbuhan, maka genom yang tepat harus dipilih adalah kloroplas atau inti bukan mitokondria, karena pada tumbuhan DNA mitokondria memiliki ukuran yang besar dan lebih sulit untuk diisolasi dan dipurifikasi. *Kedua*, ini berkaitan dengan metode mana yang harus dipilih, analisis restriksi atau sikuen DNA.

3. Tahapan dalam analisis filogenetika molekuler

Analisis filogenetika molekuler merupakan proses bertahap untuk mengolah data

sikuen DNA atau protein sehingga diperoleh suatu hasil yang menggambarkan estimasi mengenai hubungan evolusi suatu kelompok organisme. Ada sejumlah asumsi yang harus diperhatikan sebelum menggunakan data sikuen DNA atau protein ke analisis, diantaranya yaitu (1) sikuen berasal dari sumber yang spesifik, apakah dari inti, kloroplas atau mitokondria; (2) sikuen bersifat homolog (diturunkan dari satu nenek moyang); (3) sikuen memiliki sejarah evolusi yang sama (misalnya bukan dari campuran DNA inti dan mitokondria); dan (4) setiap sikuen berkembang secara bebas.

Paling sedikit, ada tiga tahap penting dalam analisis filogenetika molekuler, yaitu *sequence alignment*, rekonstruksi pohon filogenetika, dan evaluasi pohon filogenetika dengan uji statistik.

3.1 *Sequences alignment*

Tujuan utama dari tahap ini adalah untuk menentukan apakah satu sikuen DNA atau protein adalah homolog dengan yang lainnya. *Alignment* yang melibatkan dua sikuen yang homolog disebut *pairwise alignment*, sedangkan yang melibatkan banyak sikuen yang homolog disebut *multiple alignment*. Keberhasilan analisis filogenetika sangat tergantung kepada akurasi proses *alignment*. Saat ini, banyak program komputer tersedia secara gratis di internet untuk membantu proses *alignment*, misalnya ClustalX.

Dalam studi filogenetika molekuler, tahap ini akan menjadikan, misalnya setiap basa nukleotida (A,C,T,G), menjadi *site* tertentu yang ekuivalen dengan karakter, seperti karakter lebar daun atau sifat permukaan batang ketika menggunakan data morfologi. Jadi, misalnya diperoleh ukuran sikuen DNA sepanjang 600 pasang basa, maka jumlah karakter yang digunakan adalah sebanyak 600 karakter.

Dalam proses *alignment* sering ditemukan adanya *gap*, yang ditandai oleh

garis putus-putus. *Gap* terjadi karena adanya insersi dan atau delesi. Dalam prakteknya, *gap* bisa dianggap sebagai data yang hilang, walaupun dalam banyak kasus *gap* dapat dilibatkan dalam analisis karena bisa bersifat informatif.

3.2 Rekonstruksi pohon filogenetika

Metode dalam membangun suatu pohon filogenetika dengan menggunakan karakter molekuler, sikuen DNA misalnya, dibagi menjadi empat kelompok utama, yaitu *distance method* (DM), *likelihood method* (LM), *Bayesian method* (BM), dan *parsimony method* (PM). Prinsip DM adalah jumlah perbedaan nukleotida antara dua sikuen DNA menunjukkan jarak evolusi yang terjadi. Jarak evolusi dihitung untuk semua pasang sikuen DNA dan sebuah pohon filogenetika direkonstruksi dari jarak atau perbedaan pasangan basa nukleotida tersebut dengan menggunakan kriteria *least square*, *minimum evolution*, *neighbor joining*, dan *distance measure*.

Sebelum digunakan untuk merekonstruksi pohon filogenetika, LM telah lama digunakan untuk data frekuensi gen. Prinsip dari LM ini adalah bahwa perubahan-perubahan diantara semua basa nukleotida adalah sebanding. Masalah serius dari metode ini adalah waktu perhitungan yang lama, walaupun telah dikembangkan algoritma baru yang dianggap dapat mempercepat proses perhitungan.

Untuk BM, pada dasarnya adalah sama dengan LM, hanya berbeda dalam penghitungan distribusi *prior* untuk membangun pohon filogenetika. Salah satu metode untuk menghitung distribusi *prior* adalah metode MCMC (Markov chain Monte Carlo).

PM beranggapan bahwa perubahan mutasional berlangsung pada semua arah diantara empat basa nukleotida atau 20 asam amino yang berbeda dan, berbeda dengan ketiga metode yang lain, hanya jumlah perubahan basa nukleotida atau asam amino yang terkecil yang dapat memberikan penjelasan yang baik mengenai

keseluruhan proses evolusi yang terjadi. Kemudian, topologi pohon yang dipilih sebagai yang terbaik adalah yang mengalami jumlah perubahan yang paling kecil.

Dari keempat metode di atas, PM sangat sering dipilih, antara lain karena pohon yang dibentuk lebih menggambarkan perubahan evolusioner yang terjadi setiap waktu, mengandung asumsi bahwa proses evolusi akan menempuh jalan yang paling singkat (*parsimonious*), dan perhitungan relatif lebih sederhana dan cepat dengan tingkat realibilitas yang tinggi.

3.3 Evaluasi pohon filogenetika

Evaluasi pohon filogenetika berkaitan dengan uji reliabilitas dari sebuah pohon dan uji topologi antara dua atau lebih pohon yang berbeda berdasarkan set data yang sama. Banyak metode telah dikembangkan untuk menguji reliabilitas, diantaranya yaitu *interior branch test* (IB) dan *Felsenstein's bootstrap test* (FB). Prinsip IB adalah estimasi pohon dengan menguji reliabilitas setiap cabang sebelah dalam (*interior branch*).

Pada FB, reliabilitas diuji dengan menggunakan metode Efron's bootstrap. Sebuah set dari site basa nukleotida dicuplik secara acak yang dilakukan secara berulang, kemudian dilakukan konsensus, sehingga hanya satu pohon filogenetika yang dihasilkan. Karena pada dasarnya pola perubahan basa nukleotida sangat rumit dan sering berubah sejalan dengan waktu evolusi, metode FB sangat baik digunakan dalam mengevaluasi pohon filogenetika.

4. Program komputer

Saat ini terdapat dua program komputer utama yang sering digunakan untuk merekonstruksi pohon filogenetika, yaitu PAUP dan MrBayes. PAUP (*Phylogeny Analysis of Using Parsimony*) merupakan paket yang menyediakan banyak program untuk menyelesaikan berbagai aspek dalam analisis filogenetika molekuler. Paket

antara lain terdiri dari program untuk menyusun format data sikuen DNA atau protein, untuk merekonstruksi pohon filogenetika (berdasarkan metode parsimoni), dan untuk evaluasi pohon filogenetika. Perlu dicatat bahwa program PAUP dapat *me-run* set data selain molekuler, misalnya morfologi. Program PAUP dapat dijalankan baik dengan menggunakan komputer ber-OS (*operating system*) Macintosh dan Windows.

Seperti program PAUP, MrBayes merupakan program multifungsi untuk merekonstruksi pohon filogenetika (berdasarkan metode Bayesian). Sayangnya, program ini dibuat khusus hanya untuk set data molekuler. Program MrBayes dapat diinstal pada berbagai macam OS (*operating system*) komputer: Macintosh, Windows, Unix, dan Linux. Untuk memperoleh hasil yang optimum, komputer harus memiliki *speed* yang tinggi dan memori penyimpanan data yang besar.

Bahan Bacaan

- Baxevanis AD, Oulette BFE. 1998. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins. A John Wiley & Sons Inc., London
- van Welzen PC (1998) Phylogenetic versus Linnaean taxonomy, the continuing story. *Taxon* 47:413-423
- Cavalli-Sforza LL, Balfourier. 1997. Phylogenetic analysis: models and estimation procedures. *Am J Human Genet* 19:122-257
- de Queiroz K, Gauthier J (1990) Phylogeny as a central principle in taxonomy: phylogenetic definitions of taxon names. *Syst Zool* 39:307-322
- Felsenstein J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 17:68-76

- Fitch WM (1971) Toward defining the course of evolution. Minimal change for a specific tree topology. *Syst Zool* 20:406-416
- Hennig W (1966) *Phylogenetic systematics*. University of Illinois press, Urbana
- Hillis DM, Moritz C, Mable BK. 1996. *Molecular Systematic* 2nd Edition. Sinaeur Associates, Massachusetts
- Kluge AG, Farris JS (1969) Quantitative phyletics and the evolution of anurans. *Syst Zool* 18:1-32
- Li W, Graur D. 1991. *Fundamentals of Molecular Evolution*. Sinaeur Associates, Massachusetts
- Li W. 1997. *Molecular Evolution*. Sinaeur Associates, Massachusetts
- Nei M. 1996. Phylogenetic analysis in molecular genetic. *Ann Rev Genet* 30:371-403
- Page RDM, Holmes EC. 1998. *Molecular Evolution: A phylogenetic approach*. Blackwell Science Ltd., Liverpool
- Systematics Agenda 2000: Charting the Biosphere*. Technical Report. Herbarium, New York Botanical Garden, USA
- West JG, Faith DP. 1990. Data, methods and assumptions in phylogenetic inference. *Australian Syst Bot* 3:9-20