

**PENCIRIAN BERDASARKAN KOMPUTER ELEMEN LTR-
RETROTRANSPOSON Ty3-GYPSY DARI PANGKALAN DATA PROJEK
GENOM PADI**

MOHD AZWAN BIN ADNAN

**PERPUSTAKAAN
UNIVERSITI MALAYSIA SABAH**

**DISETASI INI DIKEMUKAKAN UNTUK MEMENUHI SEBAHAGIAN
DARIPADA SYARAT MEMPEROLEHI IJAZAH SARJANA MUDA SAINS
DENGAN KEPUJIAN**

**PROGRAM TEKNOLOGI TUMBUHAN
SEKOLAH SAINS DAN TEKNOLOGI
UNIVERSITI MALAYSIA SABAH**

2007

UNIVERSITI MALAYSIA SABAH

BORANG PENGESAHAN STATUS TESIS@

JUDUL: Pencirian Berkomputer Elemen LTR-RetrotransposonDi dalam Bengkela Data Genom PadiIJAZAH: Sarjana Muda Teknologi TumbuhanSAYA Mohd Azwan bin Adnan SESI PENGAJIAN: 2005 / 2006
(HURUF BESAR)

mengaku membenarkan tesis (LPSM/Sarjana/Doktor Falsafah) ini disimpan di Perpustakaan Universiti Malaysia Sabah dengan syarat-syarat kegunaan seperti berikut:-

1. Tesis adalah hak milik Universiti Malaysia Sabah.
2. Perpustakaan Universiti Malaysia Sabah dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. Sila tandakan (/)

SULIT

(Mengandungi maklumat yang berdarjah keselamatan atau Kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)

TERHAD

(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

TIDAK TERHAD

Disahkan Oleh

NURULAIN BINTI ISMAIL

LIBRARIAN

UNIVERSITI MALAYSIA SABAH

(TANDATANGAN PENULIS)

Alamat Tetap: No 25, Taman
ET 1/8

Tarikh: _____

Tarikh: 16/5/08

CATATAN:- *Potong yang tidak berkenaan.

**Jika tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa /organisasi berkenaan dengan menyatakan sekali sebab dan tempoh tesis ini perlu dikelaskan sebagai SULIT dan TERHAD.

@Tesis dimaksudkan sebagai tesis bagi Ijazah Doktor Falsafah dan Sarjana secara penyelidikan atau disertai bagi pengajian secara kerja kursus dan Laporan Projek Sarjana Muda (LPSM).



PENGAKUAN

Saya akui karya ini adalah hasil kerja kecuali nukilan dan ringkasan yang setiap satunya telah dijelaskan sumbernya.

10 April 2008



MOHD AZWAN BIN ADNAN
HS2005-1238



UMS
UNIVERSITI MALAYSIA SABAH

DIPERAKUKAN OLEH**Tandatangan****i. PENYELIA**

(DR. ROZIAH HJ. KAMBOL)

**ii. PEMERIKSA 1**

(PROF. MADYA DR. MARKUS ATONG)

**iii. DEKAN**

(SUPT/KS. PROF. MADYA DR. SHARIFF

.K OMANG, ADK)



PENGHARGAAN

Pada kesempatan ini, saya memanjatkan kesyukuran yang tidak terhingga ke hadrat Illahi kerana dengan izin dan limpah kurnia-Nya, kajian disertasi ini dapat disempurnakan dengan baik dan mengikut masa yang telah ditetapkan dalam memenuhi syarat penganugerahan Ijazah Sarjana Muda Sains Dengan Kepujian (Teknologi Tumbuhan), University Malaysia Sabah.

Jutaan terima kasih yang tidak terhingga kepada penyelia projek tahun akhir iaitu Dr. Roziah Kambol yang telah banyak membantu, memberi tunjuk ajar serta pandangan dalam menyempurnakan penulisan disertasi ini dan juga memberi kepercayaan penuh kepada saya yang juga bukan pelajar Bioteknologi untuk melaksanakan projek yang bermakna ini. Tidak lupa juga teman seperjuang saya, Mohd Faizul Azlan Bin Hilmy yang telah sama sama berusaha, berbincang dan membantu dalam menyiapkan disertasi ini. Jutaan terima kasih kepada Cik Nurul Idzyani bte Ibrahim kerana banyak memberi dorongan dan bantuan penulisan. Penghargaan juga ditujukan kepada kakitangan pejabat Sekolah Sains dan Teknologi, pensyarah –pensyarah terutamanya kepada Prof, Madya Dr, Markus Atong yang telah memberi tunjuk ajar dalam bidang Bioinformatik, Datin Mariam yang telah memberikan nasihat pada permulaan projek, tiak lupa juga buat Mr, Jupirkly James Slip sebagai pemangku dekan SPL dan juga buat Ms Chee Fong Tying sebagai penyelaras progam Teknologi Tumbuhan. Saya ucapkan jutaan terima kasih.

Lembaran penghargaan ini saya tujukan buat bapa saudara saya iaitu Mohd Shafry bin Mohd Rahim atas sokongan dan dorongan yang telah banyak diberikan. Tidak lupa juga kepada keluarga saya di Johor, Taiping dan Grik yang telah banyak membantu saya. Di atas kesempatan ini juga, saya ingin mengucapkan ribuan terima kasih kepada semua kenalan daripada Universiti Malaysia Sabah. Jasamu semua tetap ku kenang hingga ke akhir hayat.

MOHD AZWAN BIN ADNAN

2008

ABSTRAK

Dari kajian ini, satu jujukan penuh LTR-retrotransposon Ty3-gypsy telah dicirikan daripada pangkalan data genom padi. Elemen dengan nombor akses AL662989.3 terletak pada pasangan bes 112129 hingga ke pasangan bes 125033 pada nombor akses AL662989.3 dari kromosom 4 *Oryza sativa Japonica*. Jujukan ini dikenal pasti mempunyai gen LTR-retrotransposon Ty3-gypsy dengan struktur organisasi genomik seperti berikut- 5'-LTR-gag-pol-LTR-3'. Kesemua gen ini dicirikan berdasarkan produk protein dan motif terpelihara menggunakan pangkalan data domain terpelihara (CDD) dan pangkalan data domain keluarga protein (Pfam). Gen *gag* dicirikan oleh keputusan pencirian keluarga protein dengan nilai skor sebanyak 56.1 dan nilai e serendah 1.1e-15. Bahagian gen *pol* dicirikan oleh dua produk protein iaitu transkriptase berbalik dengan skor bit 122.2 dan nilai E 1.6e-33 dan integrase dengan skor bit 126.2 dan nilai E 9.2e-35 . Dikedua-dua hujung 5' dan 3', gen retroelemen ini, diapit oleh jujukan LTR sepanjang 1094 pasangan bes setiap satu dengan 100% kesamaan. Keseluruhan saiz jujukan lengkap LTR-retrotransposon Ty3-gypsy yang berada di dalam genom padi ialah 13050 pasangan bes.

ABSTRACT

From this research, a full-length sequence of element-like retrovirus was characterized from paddy genome database. The element with the accession number AL662989.3 was located at 112129 to 125033 base pairs in chromosome 4 of *Oryza sativa japonica* genome. It as follows consists of all endogenous retrovirus organization structure 5'-LTR-gag-pol-env-LTR-3'. The entire gene was characterized based on protein product and conserve motif by using conserve domain database (CDD) and protein family database (Pfam). *Gag* gene was characterized by Pfam with a bit score of 56.1 and e value of 1.1e-15. *Pol* gene was characterized by the present of two protein products such as reverse transcriptase with bits score 122.2 and E value 1.6e-33 and integrase with bits score 126.2 and E value 9.2e-35. At the 5' and 3' ends of the gene, it was flanking by repeated long terminal repeat (LTR) with the size about 1094 base pair each and with the 100% similarity. The size of complete LTR-retrotransposon Ty3-gypsy sequence that was inside the paddy genome is about 13050 base pair.

ISI KANDUNGAN

	Muka Surat
PENGAKUAN	ii
PENGESAHAN	iii
PENGHARGAAN	iv
ABSTRAK	v
ABSTRACT	vi
SENARAI KANDUNGAN	vii
SENARAI JADUAL	ix
SENARAI RAJAH	x
SENARAI SIMBOL	xii
BAB 1 PENDAHULUAN	1
1.1 Retroelemen	1
1.2 Objektif kajian	6
BAB 2 ULASAN PERPUSTAKAAN	7
2.0 Elemen Bertransposisi	7
2.1 Retroelemen	7
2.1.1 LINEs dan SINEs	9
2.1.2 Elemen LTR-Retrotranspon	10
2.2 Virus	11
2.3 Organisasi Genomik Retroelemen dan Elemen Seperti Retroviru	13
2.4 Bioinfromatik	16
2.5 Algoritma	17
2.6 Pangkalan Data Komputer	18
2.7 Projek Pangkalan Data Genom Padi	18
2.8 Program Pencarian BLAST	20
2.9 Bacaan Rangka Terbuka	21
2.10 Program Keluarga Protein Pfam	22
BAB 3 KAEDAH	32
3.1 Mendapatkan Jujukan Transkriptase Berbalik Elemen LTR-retrotranspon Ty3-gypsy Yang Bertindak Sebagai Prob.	24
3.2 Penyaringan Bank Gen: Pencirian Jujukan-Jujukan Yang Homolog	26
3.3 Penterjemahan Jujukan-Jujukan Nukleotida Kepada Protein	36
3.3.1 Program BCM Launcher	36

3.4 Penganalisaan Jujukan	39
3.4.1 Pencarian BLAST kedua	39
3.4.2 Penyusunan Pangkalan Data Keluarga Protein (Pfam)	40
3.5 Penganalisaan Jujukan LTR	41
BAB 4 KEPUTUSAN DAN ANALISIS DATA	43
4.1 Jujukan Asid Amino Yang Bertindak Sebagai Prob	43
4.2 Keputusan Pencarian BLAST	44
4.3 Keputusan Pencarian Bingkai Bacaan Terbuka Menggunakan Program ORF finder.	48
4.3.2 Keputusan Penterjemahan 3 Rangka Bingkai Bacaan Terbuka Nukleotida Kepada Protein Melalui Program BCM Search Launcher.	50
4.4 Pencarian Pangkalan Data Domain Terpelihara (CDD).	71
4.5 Pengesahan Melalui Program Pfam	73
4.6 Penentuan Jujukan LTR dengan menggunakan BLAST2seq.	74
4.7 Pencirian gen retrovirus berdasarkan pencarian domain-domain Terpelihara	76
BAB 5 PERBINCANGAN	77
5.1 Pencarian Jujukan Transkriptase Berbalik Yang Bertindak Sebagai Prob	77
5.2 Penyaringan BLAST Bagi Jujukan Yang Homolog	79
5.3 Penganalisaan LTR Melalui Program BLAST2Seq	83
5.4 Pencirian Keseluruhan Elemen LTR-retrotransposon Ty3-gypsy	84
5.4.1 Pencirian Motif-Motif Terpelihara	86
5.4.2 Pencirian Gen <i>Gag</i>	88
5.4.3 Pencirian Gen <i>pol</i>	89
BAB 6 KESIMPULAN	92
RUJUKAN	94

SENARAI JADUAL

No. Jadual	Muka Surat
2.1 Senarai Family virus jenis RNA transkrib berbalik.	12
3.1 Penerangan komponen komponen di dalam kotak pencarian BLAST	28
3.2 Jenis jenis pangkalan data	29
3.3 Jenis jenis program yang terdapat di dalam BLAST	30
4.1 Keputusan BLAST yang menunjukkan nilai skor bit, nilai jangkaan e, maklumat jujukan serta <i>hyperlink</i> untuk memasuki pangkalan data genom.	45
4.2 Jadual mmenerangkan penunjuk warna yang digunakan dalam kajian ini.	48
4.3 Kedudukan serta jenis jenis motif terpelihara yang digunakan untuk menunjukkan kehadiran sesuatu protein di dalam jujukan.	76

SENARAI RAJAH

No. Rajah	Muka Surat
2.1 Struktur gen bagi jenis-jenis retroelemen.	8
2.2 Organisasi elemen LTR-retrotransposon	14
3.1 Grafik pencarian jujukan protein transkriptase berbalik	25
3.2 Contoh prob transkriptase berbalik.	26
3.3 Kotak pencarian BLAST .	27
3.4 Contoh keputusan pencarian BLAST secara paparan grafik.	31
3.5 Contoh keputusan BLAST secara paparan senarai nilai.	32
3.6 Contoh paparan dalam bentuk susunan jujukan dan keputusan statistik	33
3.7 Paparan grafik bagi kotak pencarian bacaan rangka terbuka (ORF).	34
3.8 Contoh keputusan daripada pencarian bacaan rangka terbuka (ORF).	35
3.9 Paparan grafik bagi progam BCM <i>serach launcher</i> .	37
3.10 Paparan grafik Pfam.	40
3.11 Paparan grafik bagi progam BLAST2seq.	41
4.1 Jujukan protein Athila4 yang digunakan sebagai prob	43
4.2 Keputusan BLAST melalui paparan grafik	44
4.3 Keputusan Blast yang menunjukkan penjajaran bagi elemen Athila dengan jujukan yang homolog pada pangkalan data genom.	47
4.4 Keputusan Paparan Grafik bagi Program ORF finder	48
4.5 Keputusan BLAST kedua daripada program ORF finder melalui Program Pangkalan Data Domain Terpelihara (CDD).	49
4.6 Produk Protein: Gag	71
4.7 Produk Protein :Transkriptase Berbalik	71
4.8 Produk Protein: Transkriptase Berbalik	72
4.9 Produk Protein: Integrase	72
4.10 Keputusan paparan grafik bagi pencarian domain-domain terpelihara pada pangkalan data protein Pfam.	73

4.11 Keputusan pencarian protein berfungsi Pfam pada rangka ketiga penterjemahan nukleotida.	74
4.12 Penjajaran pada kedua dua hujung LTR.	74
4.13 Keputusan dari penjajaran dalam bentuk grafik.	75

SENARAI SIMBOL

ALV	Avian Leukosis Virus
ALSV	Avian Leukosis-Sarcoma Virus
BLAST	Basic Local Alignment Search Tool
BLV	Bovine Leukemia Virus
DNA	Deoxyribonuclease acid
EAV	Avian Endogenous Retrovirus
<i>env</i>	Envelope
EST	Expressed Sequence Tagged
E Value	Expectation Value
<i>gag</i>	Group Specific Antigen
GSS	Genome Survey Sequences
HSV	Human Spumavirus
LTR	Long Terminal Repeat
MPMV	Mason-Pfizer Monkey Virus
NCBI	National Centre Biotechnology Information
nr	non-redundant
ORF	Open Reading Frame
RNA	Ribonucleic acid
<i>pol</i>	Polymerase
STS	Sequence Tagged Sites

WDSV	Walleye Dermal Sarcoma Virus
WGS	Whole Genome Sequence
Pfam	Protein Family
LINE	Long Intersped Nuclear Element
SINE	Short Intersped Nuclear Element

BAB 1

PENDAHULUAN

1.1 Retroelemen

Retroelemen atau dikenali sebagai elemen bertransposisi RNA, bertransposisi melalui enzim transkriptase berbalik menggunakan RNA sebagai perantaraan. Proses ini berlaku dengan retroelemen DNA ditukarkan kepada RNA kemudian ia ditukarkan kembali kepada DNA dimana ia akan berintegrasi dengan genom. Perbezaan retroelemen dengan elemen bertransposisi DNA ialah elemen ini tidak dikeluarkan dari genom semasa ia bertransposisi.

Terdapat 5 jenis retroelemen iaitu SINEs, LINEs, LTR-retrotranspon dari kumpulan Ty1-gypsy dan kumpulan Ty3-copia dan yang terakhir sekali ialah elemen retrovirus.. Secara umumnya, elemen bertransposisi terdiri daripada 2 kumpulan dan retroelemen ialah kelas pertama elemen bertransposisi dan kelas kedua ialah elemen bertransposisi DNA. Kedua-dua kelas terdapat dalam tumbuhan malah kelas I retrotranspon terdapat dalam bilangan yang banyak di dalam genomik DNA tumbuhan dari 15% di dalam *Arabidopsis thaliana* sehingga 70% hingga 80% di dalam tumbuhan

jagung atau barli (Marco dan Marin, 2005). Retrotransposon secara langsung bertanggungjawab dalam membentuk genom tumbuhan dalam bentuk mutasi gen dan menghasilkan pengubahsuain gen melalui rekombinasi (Kumar dan Bennetzen, 1999).

Retroelemen yang paling banyak terdapat dalam tumbuhan ialah LTR-retrotransposon yang mengandungi LTR dalam elemen tersebut. Ia terdiri daripada beberapa ratus pasangan bes sehingga 10 kb. Kebolehan istimewa yang terdapat elemen ini ialah pada bahagian gen integrase dimana kedudukan gen integrase membezakan LTR-retrotransposon kepada dua jenis iaitu gypsy dan copia. Enzim ini bertindak memasukkan produk hasil daripada enzim transkriptas berbalik ke dalam kromosom.

Kajian terbaru daripada beberapa ahli penyelidik mengatakan retroelemen yang mengandungi gen *env* bukan sahaja terdapat pada haiwan malah wujud dengan banyaknya pada tumbuhan (Kumar dan Bennetzen, 1999). Kajian terhadap *Arabidopsis thaliana*, *Pisum sativum*, *Glycine max*, *Gossypium barbadense* dan *Hordeum vulgare* menunjukkan kehadiran gen *env* pada genom mereka (Peterson-Burch et al, 2000).

Hasil kajian Thomas Eickbush menunjukkan retrovirus dan LTR-retrotransposons secara evolusinya adalah berkait rapat antara satu sama lain. Dalam kajian enzim transkriptas berbalik bagi retrovirus yang dilakukan oleh Xiong dan Eickbush (1990), mereka mendapati enzim transkriptase berbalik bagi retrovirus sangat menyerupai enzim transkriptase berbalik bagi LTR-retrotransposon. Ini menyebabkan spekulasi tentang

kewujudan retrovirus, dimana ia dikatakan berasal daripada LTR retrotransposon (Marco dan Marin, 2005).

ORF ketiga yang mengandungi gen *env* telah pun dijumpai di dalam tumbuhan dan kali pertama dijumpai oleh Wright dan Voytas (1998). Mereka menerangkan bahawa beberapa elemen *Arabidopsis thaliana* yang di gelar Athila mengandungi gen-gen seperti retrovirus malah mengandungi ORF ketiga yang mengekodkan protein *env*. Malah dalam analisa filogenetik yang mereka lakukan, elemen Athila mempunyai hubungan yang sangat rapat dengan elemen Ty3/gypsy tumbuhan yang lain. Ini menunjukkan Athila adalah merupakan sejenis elemen LTR-retrotransposon dan dalam analisa filogeni mempunyai persamaan yang jauh dengan errantivirus serangga (Wright dan Voytas, 1998). Kajian ini membuktikan kehadiran elemen seperti retrovirus di dalam tumbuhan dan perbezaannya dengan elemen Athila tersebut terhadap elemen kumpulan Ty3-gypsy serangga melalui analisa evolusi.

Kajian-kajian terhadap elemen errantivirus ini diteruskan lagi dengan mengkaji lebih mendalam lagi tumbuhan *Arabidopsis* dan seterusnya tumbuhan tumbuhan lain. Elemen seperti Athila kecil, cyclops, calypso, SIRE-1 merupakan elemen retrotransposon yang mempunyai gen-gen yang sama seperti retrovirus (Kumar dan Bennetzen, 1999) . Kajian terdahulu hanya mengaitkan elemen dalam kumpulan Ty1 gypsy sebagai elemen seperti retrovirus. Kajian pada tumbuhan kacang soya memberikan hasil penemuan bagi satu elemen dalam kumpulan Ty3-copia yang dipanggil SIRE-1 dimana ia mengandungi jujukan *env* (Laten et al., 1998). Kajian ini menunjukkan kehadiran gen *env* tidak terhad

pada LTR-retrotransposon kumpulan gypsy malah ia wujud juga di dalam kumpulan elemen copia.

Sel tumbuhan mempunyai satu struktur yang khusus iaitu dinding sel. Dinding sel tumbuhan terdiri daripada selulosa dan bahan-bahan lain yang menyebabkan struktur sel itu menjadi tetap. Retrovirus pula adalah sejenis virus selaput membran iaitu ia memasuki sel perumah secara menembusi bahagian dwilapis membran. Halangan utama bagi retrovirus menyerang sel tumbuhan ialah dinding sel tumbuhan dan menyebabkan satu spekulasi tentang kewujudan retrovirus itu sendiri di dalam tumbuhan. Terdapat beberapa andaian dibuat. Andaian pertama ialah retrovirus memasuki sel tumbuhan melalui liang plasmodesmata (Kumar dan Bennetzen, 1999). Jika andaian ini benar maka peranan gen *env* pada tumbuhan adalah berlainan dengan peranan gen *env* pada retrovirus. Andaian kedua pula ialah kitar hidup bagi retrovirus bagi tumbuhan adalah sama dengan kitar hidup retrovirus bagi haiwan dimana ia berpindah dari perumah haiwan kepada perumah tumbuhan (Laten, et al., 1998). Retrotransposon terdapat pada semua jenis tumbuhan seperti gymnosperma dan angiosperma. Tetapi kajian lebih tertumpu kepada tumbuhan monokotiledon dan dikotiledon. Elemen-elemen LTR retrotransposon yang mengandungi gen *env* kini telah dijumpai pada kedua-dua kumpulan tumbuhan tersebut.

Kajian ini dilakukan untuk mencirikan struktur retroelementen jenis LTR-retrotransposon Ty3-gypsy di dalam genom padi. Elemen LTR-retrotransposon Ty3-gypsy dicirikan dengan menggunakan kaedah-kaedah bioinformatik. Terdapat kajian-kajian terdahulu yang mengkaji retroelementen di dalam padi (McCarthy et al., 2002). Maka

kajian ini khusus kepada mencirikan elemen LTR-retrotransposon Ty3-gypsy di dalam pangkalan data genom padi.

Pencirian bagi LTR-retrotransposon Ty3-gypsy dapat membantu ahli penyelidik pada masa akan datang memahami lebih lanjut lagi tentang retroelement tumbuhan dan mungkin dapat memberikan lebih gambaran tentang evolusi retroelement tumbuhan. Kini, retrovirus memainkan peranan penting sebagai vektor untuk memindahkan gen dalam haiwan. Mungkin pada masa akan datang, vektor retrovirus dapat digunakan dalam tumbuhan dengan memahami dengan secara terperinci lagi struktur dan peranan elemen seperti retrovirus di dalam tumbuhan. Retroelement kini digunakan dengan meluas sebagai penanda molekular di dalam makmal (Bennetzen, 2000). Peranan retroelement di dalam organisma masih kurang jelas melainkan ia dikatakan sebagai gen sampah namun ia bertanggungjawab dalam membentuk struktur genetik bagi sesuatu organisma terutamanya tumbuhan (Bennetzen, 2000).

1.2 OBJEKTIF KAJIAN

Objektif kajian ini bertujuan:

- 1) Mendapatkan dan mengenalpasti prob yang sesuai untuk menyaring LTR-retrotransposon Ty3-gypsy di dalam pangkalan data genom padi.
- 2) Menyaring jujukan yang homolog pada pangkalan data padi(*Oryza sativa Japonica*) dengan menggunakan kaedah BLAST pada laman web NCBI.
- 3) Menterjemahkan jujukan-jujukan nukleotida kepada jujukan-jujukan asid amino dengan menggunakan program BCM *search launcher* atau ORF *finder*.
- 4) Mencirikan bahagian-bahagian gen yang terdapat pada elemen-seperi retrovirus dari jujukan yang diperoleh daripada pangkalan data projek padi.

BAB 2

ULASAN PERPUSTAKAAN

2.0 Elemen Bertransposisi

Elemen bertransposisi (TEs) atau dikenali sebagai transposon merupakan elemen genetik yang dapat mengubah kedudukan atau bergerak dalam genom. Ia juga dikenali sebagai gen lompat di mana ia boleh berintegrasi pada bahagian-bahagian lain dalam genom secara bergerak dari satu bahagian ke bahagian yang lain iaitu secara menghasilkan salinan genetik mereka sendiri kemudian dimasukkan dalam bahagian genom organisma tertentu. Secara amnya, terdapat dua kelas utama bagi elemen ini iaitu elemen bertransposisi DNA dan retroelemen. (Bennetzen, 2000)

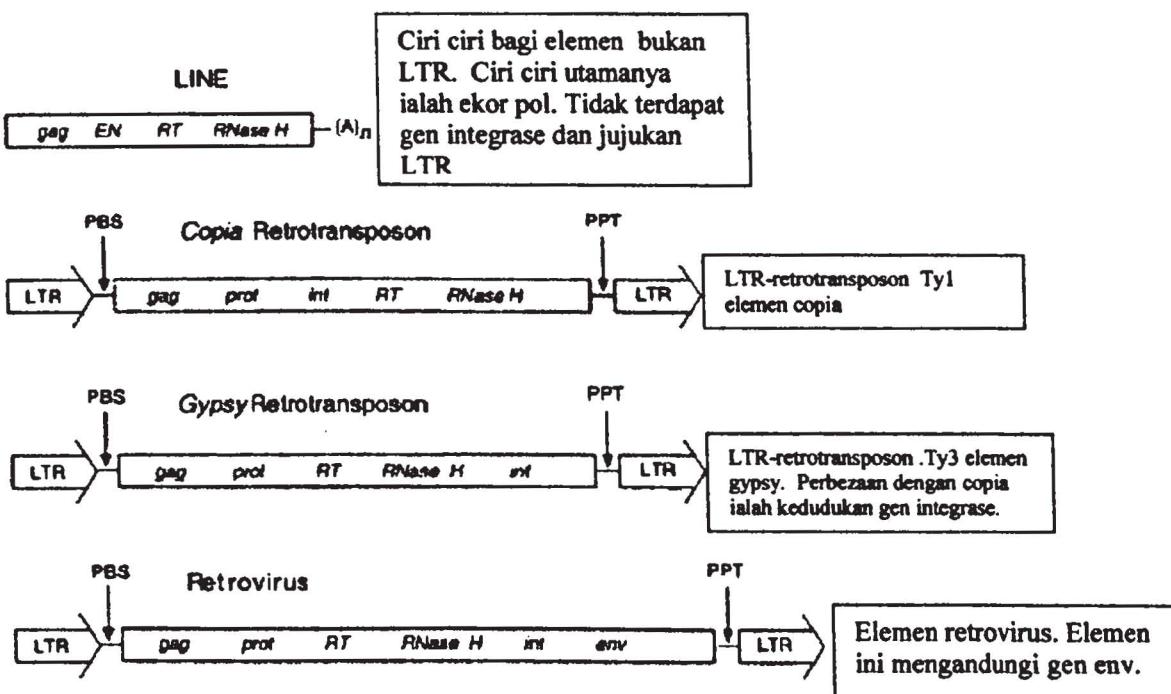
2.1 Retroelemen

Retroelemen mendapat namanya daripada gen transkriptase berbalik iaitu perkataan retro membawa maksud berpatah balik dalam bahasa Latin. Retroelemen merupakan salah satu kumpulan utama daripada elemen transposisi. Ia merupakan komponen yang banyak terdapat dalam organisma eukaryotik dan prokaryotik. Ia merangkumi pelbagai jenis

elemen genetik seperti virus RNA dan DNA, retron bakteria, kumpulan kedua intron dan retrotransposon (Laten, H et al, 1998).

Elemen retrovirus merupakan retroelement dan seperti retroelement yang lain, retrovirus menggunakan enzim transkriptase berbalik untuk menukar RNA kepada DNA. Rajah 2.1 di bawah menunjukkan struktur gen bagi jenis-jenis retroelement seperti LINEs, SINEs, LTR-retrotransposon dan retrovirus (Bennetzen, 2000).

Rajah 2.1 Rajah menunjukkan struktur gen bagi jenis-jenis retroelement.



2.1.1 LINEs dan SINEs

LINEs (long intersped nuclear elements) merupakan salah satu cabang dari retroelement jenis retrotransposon bukan LTR di mana seperti rajah 2.1, tidak terdapat jujukan LTR pada hujung 5' dan 3' permulaan dan penghujung elemen itu. Ia mengandungi ekor poly(A). LINEs mempunyai panjang 1 hingga 2 ribu pasangan bas dan mengandungi 2 ORF yang mempunyai bahagian *pol* yang mengekodkan protein enzim transkriptase berbalik serta enzim endonuklease.

Kedua-dua bahagian *pol* tersebut bertanggungjawab bagi membolehkan elemen ini bertransposisi. Perbezaan LINEs dengan LTR-retrotransposon ialah ekor poly (A) dan jujukan LTR pada elemen tersebut. Poly (A) berfungsi menukarkan RNA LINEs kepada DNA dan endonuklease berfungsi memasukkan DNA LINEs itu ke dalam genom tertentu (Schmidt, 1999). Satu lagi bahagian dalam LTR bukan retrotransposon ialah SINEs (short intersped nuclear elements) . Ia terdiri daripada beberapa ratus pasangan bas. Ia juga mengandungi ekor poly(A). SINEs tidak mengekodkan enzim transkriptase berbalik dan tidak dapat bergerak secara terus. Namun, seperti juga LINEs, ia bergerak secara retrotransposisi (Schmidt, 1999) .

2.1.2 LTR-Retrotransposon

Retroelemen yang paling banyak terdapat dalam tumbuhan ialah LTR-retrotransposon yang mengandungi LTR (long terminal repeat) dalam elemen tersebut. Keseluruhan genom LTR-retrotransposon terdiri daripada beberapa 10 ribu pasangan bes sehingga 14 ribu bes. Kebolehan istimewa yang terdapat pada elemen ini ialah pada bahagian gen integrase di mana kedudukan gen integrase membezakan elemen LTR-retrotransposon kepada dua kumpulan iaitu gypsy dan copia. Bahagian gen integrase bagi elemen copia terletak di antara kedudukan gen transkriptase berbalik dan gen protease manakala bagi elemen gypsy, gen integrase terletak pada bahagian akhir elemen iaitu selepas gen RNaseH. Kedudukan gen-gen pada elemen gypsy adalah sama seperti kedudukan gen pada endogenous retrovirus. Enzim ini bertindak memasukkan produk hasil daripada enzim transkriptase berbalik ke dalam kromosom. Dalam haiwan, retrotransposon jenis gypsy mengandungi satu bahagian gen yang mengkodkan protein sampul yang membenarkan ia dibungkus ke dalam lapisan membran mengakibatkan jangkitan intersetular. Retroelemen ini digelar retrovirus dan pada masa sekarang dikatakan hanya terdapat pada haiwan (Kumar dan Bennetzen, 1999).

RUJUKAN

Abdel Ghany A.A., Zaki E.A., 2002. Cloning and sequencing of an envelope-like gene in *Gossypium*. *Planta* 216:351-353.

Bennetzen, J.L., 2000. Transposable elements contributions to plant gene dan genome evolution. *Plant Mol. Biol.* 42:351-369.

Doolittle, R.F., 1990. Retrovirus phylogeny and evolution. *Curr. Top. Microbiol. Immunol.* 157: 1 – 18.

Emilio, A., Emini, 2002. *The Human Immunodeficiency Virus*, Princeton University Press, Amerika Syarikat

Essam, A. Zaki, 2003. Plant retroviruses: structure, evolution dan future applications. *African Journal of Biotechnology* Vol. 2 (6), pp. 136–139, June 2003ISSN 1684-5315 © 2003 Academic Journals.

Flint S. J, Enquist L.W , Krug R.M , Racaniello V.R, Skalka A.M 2000. Principles of Virology: *Molecular biology, pathogenesis and control*. ASM Press

Franke E.K., Yuan,H.E. and Luban,J, 1994. Specific incorporation of cyclophilin A into HIV-1 virions. *Nature*, 372:, 359–36

- Goff, S.A., Ricke, D., Lan, T.H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., 2002: A draft sequence of the rice genome (*Oryza sativa L. ssp. japonica*). *Science*, 296:92-100.
- Gogg, S. P., 1990. Retroviral reverse transcriptase : synthesis, structure, and function. *J. Acquired Immune Defic. Syndr.* 3:817-831
- Hirochika, H., Sugimoto, K., Otsuki, Y., Tsugawa, H., Kanda, M., 1996. Retrotransposons of rice involved in mutations induced by tissue culture. *Proc Natl Acad Sci USA.*;93:7783–7788.
- Claverie, J.M., dan Notredame, C, 2003. Bioinformatics for Dummies. Wiley Publication, Inc.
- Kambol, R., 2003. *Distribution And Evolution Of Endogenous Retrovirus Within Amphibian And Piscine Hosts*. Unpublished PhD. Thesis Department of Biological Science, Imperial College, London, United Kingdom.
- Kim, J.M., Vanguri, S., Boeke, J.D., Gabriel, A., Voytas, D.F., 1998. Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by the complete *Saccharomyces cerevisiae* genome sequence. *Genome Res.*;8:464–78.

Kumar, A. dan Bennetzen, J.L., 1999 . Plant retrotransposons. *Annu Rev Genet.* 1999;33:479–532.

Laten, H., Majumdar, A., dan Gaucher, E.A. 1998. SIRE-1, a copia/Ty1-like retroelement from soybean, encodes a retroviral envelope-like protein. *Proc. Natl. Acad. Sci.* 95: 6897-6902

Lerat, E. dan Capy, P. 1999. Retrotransposons dan retroviruses: Analysis of the envelope gene. *Mol. Biol. Evol.* 16: 1198-120

Lewin, 2004. Genes VIII. Pearson Education Inc. Amerika Syarikat.

Marco, A. dan Marin, I., 2005. Retrovirus-like elements in plants. *Recent Res. Devel. Plant Sci.* 3(2005): ISBN: 81-7736-245-3

McCarthy, E.M., Liu, J.D., Gao, L.Z., dan McDonald, J.F., 2002. LTR retrotransposons of *Oryza sativa*. *Genome Biol.*;3:research0053.1–0053.11. doi: 10.1186/gb-2002-3-10-research0053.

Peterson-Burch, B.D., Wright D.A., Laten, H.M., dan Voytas, D.F. 2000. Retroviruses in plants? *Trends Genet.* 16:151-152 pubmed.

Sasaki, T., Burr B., 2000. International rice genome sequencing project: The effort to completely sequence the rice genome. *Curr Opin Plant Biol* 2000, 3:138-141.

Song, S.U, Gerasimova, M, Kurkulos M, Boeke, J.D, Corces, V.C., 1994. An *env*-like protein encoded by a *Drosophila* retroelement: evidence that *gypsy* is an infectious retrovirus. *Genes Dev.* 8:2046-2057.

Vicient, C.M., Kalendar, R., Schulman, A.H. 2001. Envelope-class retrovirus-like elements are widespread, transcribed dan spliced, dan insertionally polymorphic in plants. *Genome Res.* 11:2041-2049.

William, S. Klug dan Michael. R. Cummings, 2006 Concepts Of Genetics. Pearson Prentice Hall.

Wills, J.W., dan R.C. Craven. 1991. Form, function, and use of retroviral Gag proteins. *AIDS* 5:639-634

Wilhelm, M. and Wilhelm, F.X. 2001. Reverse transcription of retroviruses and LTR retrotransposons. *Cell. Mol. Life Sci.* 58:1246-1262.

Wright, D.A. dan Voytas, D.F. 1998. Potential retroviruses in plants: Tat1 is related to a group of *Arabidopsis thaliana* Ty3/gypsy retrotransposons that encode envelope-like proteins. *Genetics* 149: 703-715

Wright, D.A. dan Voytas, D.F. 2002, Athila4 of Arabidopsis dan Calypso of Soyabean Define a lineage of endogenous Plant Retroviruses. *Genome Res.* 12:122-131.

Xiong, Y., Eickbush, T.H., 1990. Origin dan evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J.* 1990;9:3353–3362.

Yu, J., Hu, S., Wang, J., Wong G.K., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang , X., Cao, M., Liu, J., Sun, J., Tang, J., Chen, Y., Huang, X., Lin, W., Ye, C., Tong, W., Cong, L., Geng, J., Han, Y., Li, L., Li, W., Hu, G., Huang, X., Li, W., Li, J., Liu, Z., Li, L., Liu, J., Qi, Q., Liu, J., Li, L., Li, T., Wang, X., Lu, H., Wu, T., Zhu, M., Ni, P., Han, H., Dong, W., Ren, X., Feng, X., Cui, P., Li, X., Wang, H., Xu, X., Zhai, W., Xu, Z., Zhang, J., He, S., Zhang, J., Xu, J., Zhang, K., Zheng, X., Dong, J., Zeng, W., Tao, L., Ye, J., Tan, J., Ren, X., Chen, X., He, J., Liu, D., Tian, W., Tian, C., Xia, H., Bao, Q., Li, G., Gao, H., Cao, T., Wang, J., Zhao, W., Li, P., Chen, W., Wang, X., Zhang, Y., Hu, J., Wang, J., Liu, S., Yang, J., Zhang, G., Xiong, Y., Li, Z., Mao, L., Zhou, C., Zhu, Z., Chen, R., Hao, B., Zheng, W., Chen, S., Guo, W., Li, G., Liu,



S, Tao, M., Wang, J., Zhu, L., Yuan, L., Yang, H.: A Draft Sequence of the Rice Genome (*Oryza sativa* L. ssp. *indica*). *Science* 2002, 296:79-92.

LAMAN WEB :

1) Pangkalan data genom:

Pangkalan Data Genom Padi

http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?taxid=4530

2) Pangkalan Data Protein

Perisian Famili Protein (Pfam)

<http://pfam.sanger.ac.uk/>

Perisian Pengkalan Data Domain Terpelihara (CDD)

<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>

3) Analisis Protein dan Nukleotida

Perisian Anotasi Protein

<http://searchlauncher.bcm.tmc.edu/seq-util/seq-util.html>

Perisian Bacaan Rangka Terbuka (ORF)

<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>



UMS
UNIVERSITI MALAYSIA SABAH

Perisian BLAST

<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>

Perisian Pencarian LTR

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi>