

台灣生物資訊核心設施

Taiwan Bioinformatics Institute Core Facility

I-Shou Chang, Ping-Chiang Lyu
Chao A. Hsiung, Hsien-Da Huang, H. Sunny Sun

National Health Research Institutes, National Tsing Hua University,
National Chiao Tung University, National Cheng Kung University

TBI Core Facility

NGS data analysis

Statistical Genetics

Computational proteomics

Comparative Genomics and Interactomes

Structural Bioinformatics

Applied Medical Genomics

Missions:

Services

R & D

Education and Training

Collaborative Research

Supporting NRPB and Life Sciences

Cancer research

Infectious diseases

Highly heritable diseases

Cardiovascular diseases and Metabolic Syndrome

Neuronal Diseases

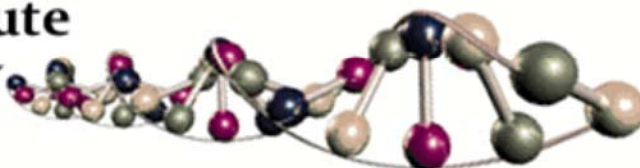
Translational Medicine

NRPB Resources Center



台灣生物資訊核心設施

Taiwan Bioinformatics Institute Core Facility

[About Us](#)[Tools](#)[Databases](#)[Education](#)[Web Resources](#)[Newsletter](#)[Help](#)[Coordination](#)[Statistical gen](#)[Computational](#)[Comparative g](#)[Structural Bio](#)[Applied Medic](#)[New Version](#)[GCG Applica](#)[Homology &](#)[Sequence An](#)[Phylogenetic Analysis](#)[Microarray Analysis](#)[Structural Analysis](#)[Miscellaneous Tools](#)[Tools Download](#)[Nucleotide D](#)[Protein Data](#)[Structure Dat](#)[FTP & others](#)[Symposium](#)[Workshop and Special Topics](#)[Tutorials](#)[On-line Resources](#)[Video Programs](#)[Course](#)

「基因體分析研
電腦教室」舉辦，

「結構生物資訊」
2011.04.09

3/26 於「國家衛生研究院圖資大樓5樓
參考網頁資訊！(公佈日期 2011/08/17)

內容請參考網頁資訊！(公佈日期

[NHRI News](#)[NTHU News](#)

❖ Databases

[CAPIH](#)

Comparative Analysis of Protein Interactions for HIV-1 [NHRI]

[CNVdb](#)

A database of copy number variations across vertebrate



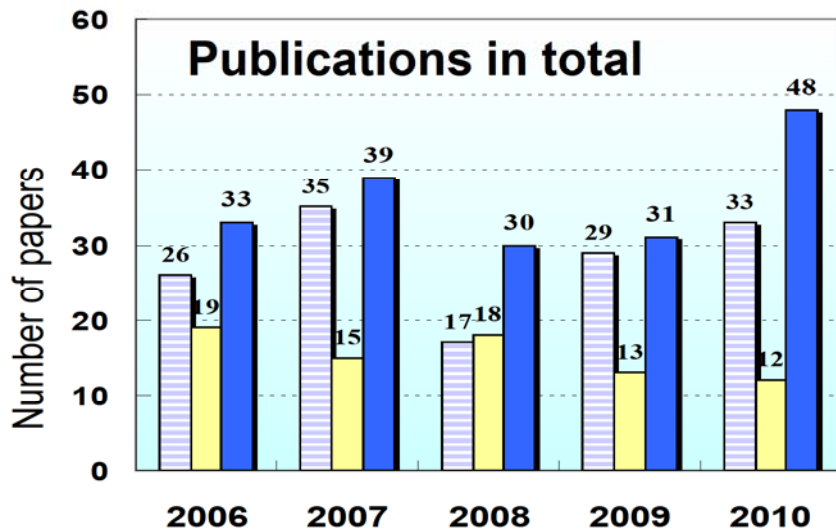
<http://www.tbi.org.tw/>

UNLab
Molecular Genetics & Bioinformatics

TBI 服務內容

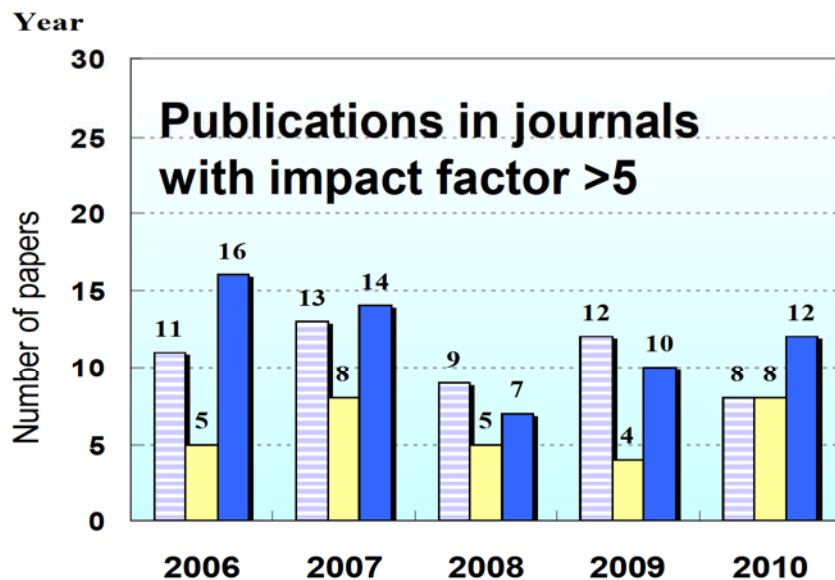
- 線上服務
 1. 40種分析工具、21種資料庫、1種鏡相站
 2. GCG, EMBOSS線上序列分析服務
- 客製化服務
- 諮詢服務

Performances of E2 Core and Bioinformatics Centers in Asia

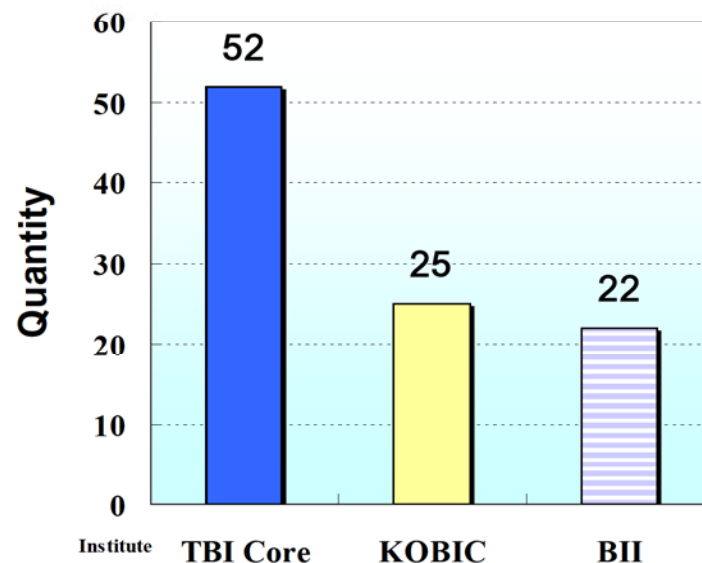


E2 Core performs better than or equally well with bioinformatics center in Korea (KOBIC) and in Singapore (BII) on publication and web tools developments.

■ E2 Core
 ■ KOBIC, Korea
 ▨ BII, Singapore



No. of web tools developed



如何使用 *NCKU online services*

網址：<http://www.binfo.ncku.edu.tw>

研究資源 → 應用資源 → 點選連結



首頁 (Home) | 簡介 (Introduction) | 本中心服務 (Service) | 研究資源 (Resources) | 相關連結 (Links) | 意見留言 (Contact)

◎應用資源

Vector NTI
IBM軟體下載
NCKU BLAST
NCKU SNP Database
Repeat Masker Server
TAG Database

成功大學-前瞻生物
成功大學生物科
討會，邀請多位名人演講，歡迎踴躍報名參加！...~ more

Gene-Based Repetitive Element Database (GBRED)

國立成功大學生物資訊研究所：梁凱強

國立成功大學分子醫學研究所：孫孝芳

What is the HGP?



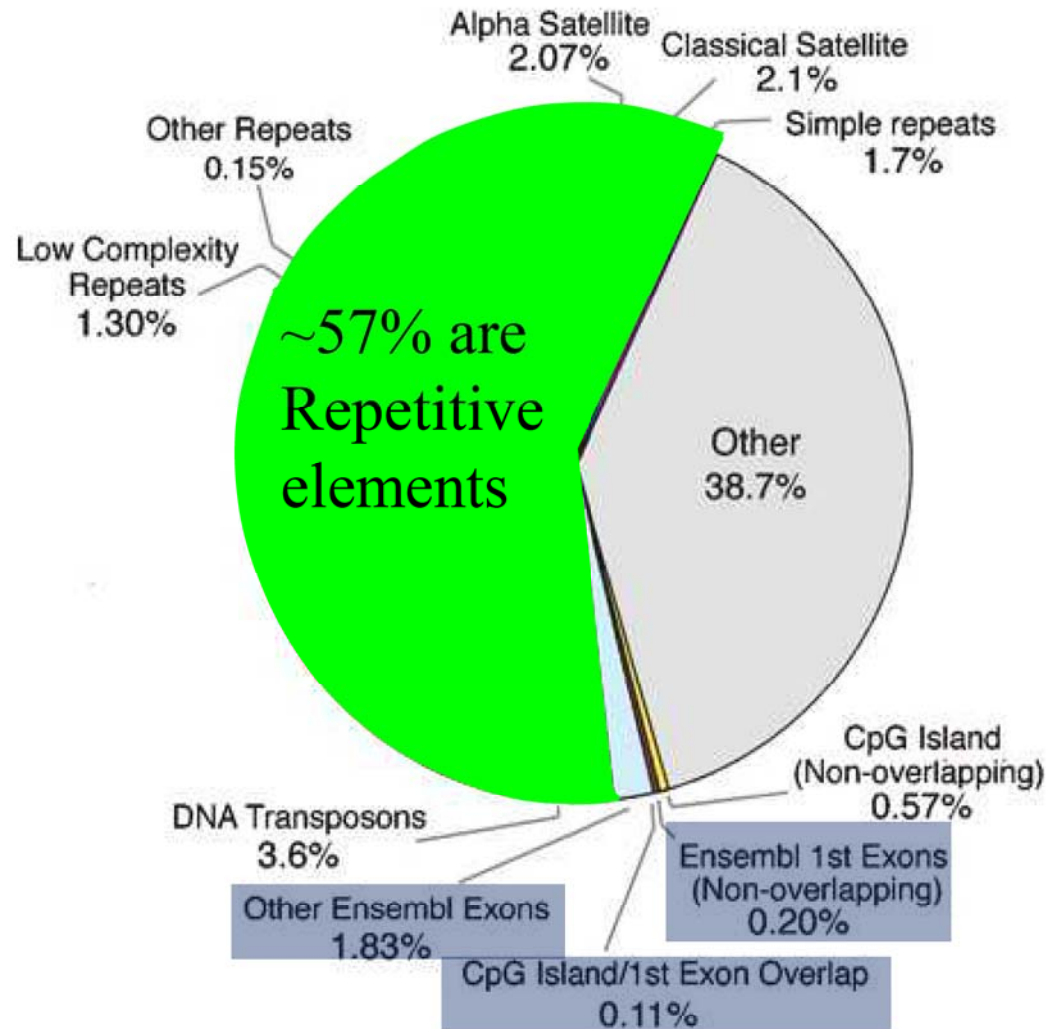
What We've Learned Thus Far

“the more we learn about the human genome, the more there is to explore.”

Repetitive elements

Class	Length	Copy Number in Human Genome
Repetitious DNA		
Simple-sequence DNA	1-500 bp	Variable
Interspersed repeats		
DNA transposons	2-3 kp	300,000
LTR retrotransposons	6-11 kp	440,000
Non-LTR retrotransposons		
LINEs	6-8 kp	860,000
SINEs	100-300 bp	1,600,000

Human genome composition



Robert A. *et al.*, *Genome Res.* 16: 157-163 (2006)

Functional roles of repetitive elements

- Repetitive elements are involved in genome rearrangement and illegitimate recombination

Kreahling *et al.*, *Trends Genet.*, 20, 1-4 (2004)

- Serve as regulatory elements controlling the transcription of nearby genes

- By acting as enhancers or promoters

Kondo, Y. *et al.*, *J. Biol. Chem.* 278: 27658-27662 (2003)

- Provide alternative splicing sites and polyadenylate tail

Makałowski, *Gene*. 259: 61-67 (2000)

- Via epigenetic mechanism

Eriko A. *et al.*, *Nucleic Acids Research*. 35: 3383-3390 (2007)

- By carry the transcription factor binding site

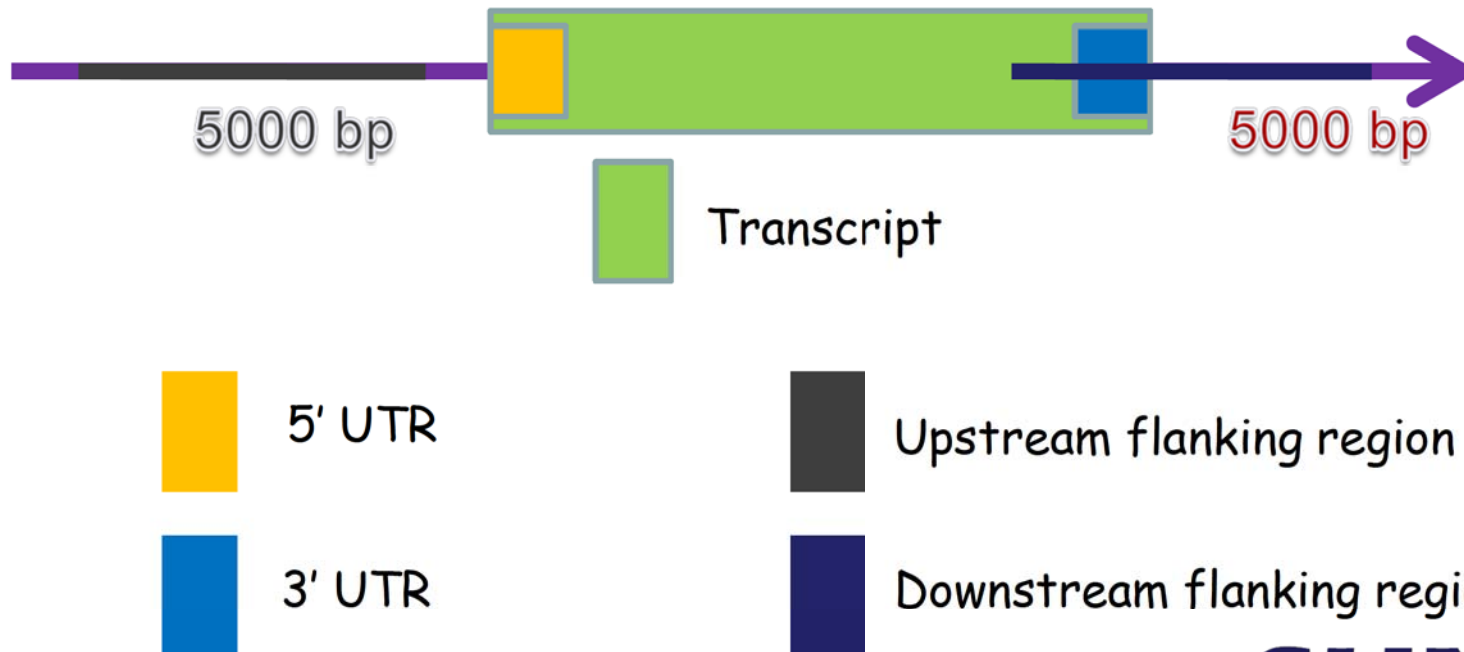
Ting W. *et al.*, *Proc Natl Acad Sci USA*. 104: 18613-18618 (2007)

Objective

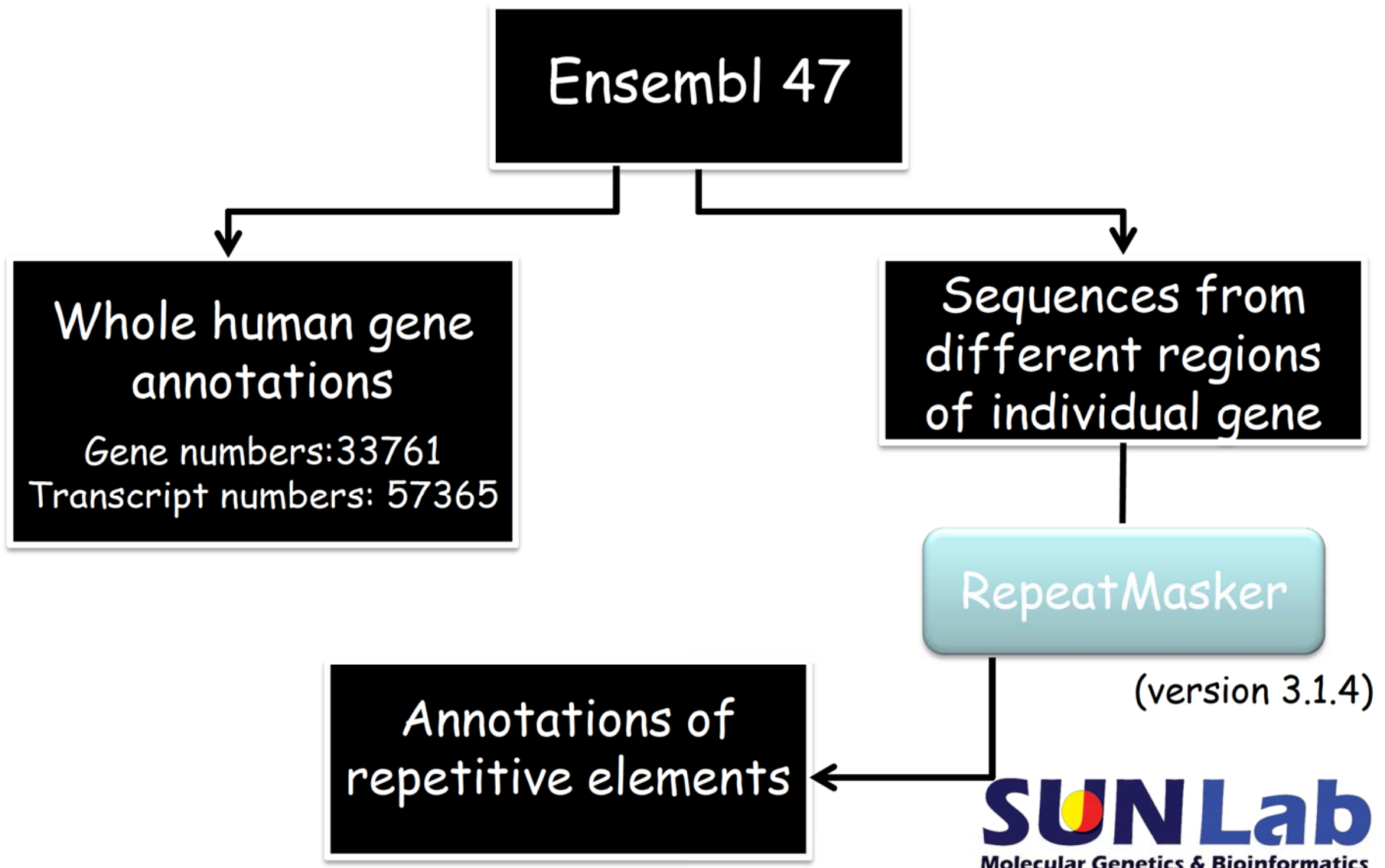
To construct a gene-based repetitive element database to provide all annotated repetitive element information that are associated with particular gene(s)

Data sources

- All sequences and annotated information of human genes were downloaded from Ensembl build version 47



Database construction



NCKU GBRED 首頁

<http://www.binfo.ncku.edu.tw/GBRED/>

The screenshot shows the NCKU GBRED website. At the top, there is a banner with the Binfo logo and the text 'NCKU Bioinformatics Center' and '國立成功大學生物資訊中心'. Below the banner is the large 'GBRED' title in blue and red, followed by 'Gene-Based Repetitive Element Database'. The main content area is divided into two columns. The left column contains links: 'GBRED', 'Homepage', 'Query', 'Manual', 'RepeatMasker', 'CONTACT US', and 'VERSION'. The right column contains an 'INTRODUCTION' section with text about the database. Callout boxes with arrows point to specific parts of the website: '進入搜尋頁面' points to the 'Query' link; 'GBRED 簡介' points to the 'INTRODUCTION' section; '進入說明頁面' points to the 'Manual' link; '進入成大 RepeatMasker 伺服器' points to the 'RepeatMasker' link; '版本及更新日期' points to the 'VERSION' section; and '寄信給成大生物資訊中心' points to the 'CONTACT US' section.

Binfo
NCKU Bioinformatics Center
國立成功大學生物資訊中心

GBRED

Gene-Based Repetitive Element Database

September 22, 2008

GBRED

Homepage
Query
Manual
RepeatMasker

CONTACT US

■ E-mail

VERSION

RepeatMasker version: 3.1.4
Rebase version : 20070801
Last updated : 20080928

INTRODUCTION

GBRED (Gene-Based Repetitive Element Database) is a dedicated database containing repetitive element information flanking all human genes, such as 5' untranslated region (5' UTR), and 3' untranslated region (3' UTR). Recent studies imply that repetitive elements which located in these regions could regulate gene expression. However, there is no systematical survey to analyze which kind of repetitive elements prefer to locate in these regions and what is the role about these repetitive elements. GBRED is trying to provide this information to users. You can click the query button in left panel to enter the information query system.

進入搜尋頁面

GBRED 簡介

進入說明頁面

版本及更新日期

進入成大
RepeatMasker 伺服器

寄信給成大生物資訊中心

GBRED 搜尋介面

選擇搜尋特定區域的重複性序列

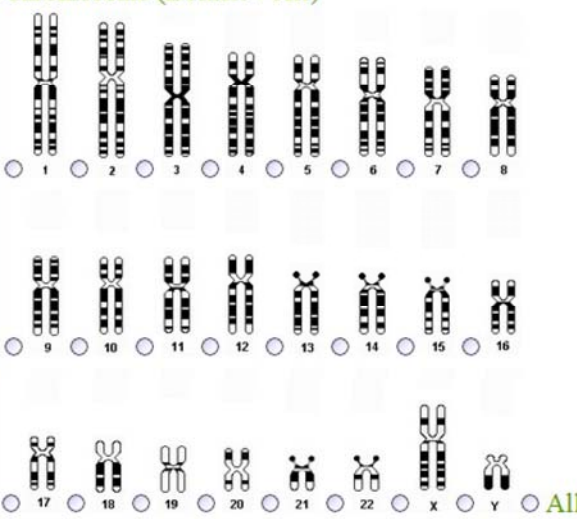
GBRED
[Homepage](#)
[Query](#)

- Manual
- RepeatMasker

CONTACT US

- E-mail

VERSION
RepeatMasker version: 3.1.4
Repbse version : 20070801
Last updated : 20080928

QUERY
Region :
RepeatClass :
Subclass :
Gene Symbol :
Chromosome (Default : All)


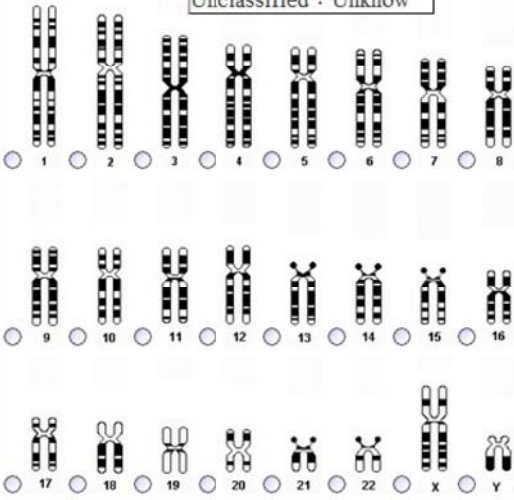
UTR means untranslated region;
Flank_u5000 means 5kb upstream
from transcription start site;
Flank_d5000 means 5kb downstream
from transcription end site.

GBRED 搜尋介面

GBRED
[Homepage](#)
[Query](#)
■ [Manual](#)
■ [RepeatMasker](#)
CONTACT US
■ [E-mail](#)
VERSION
RepeatMasker version: 3.1.4
Rebase version : 20070801
Last updated : 20080928

QUERY September 30, 2008

Region :

RepeatClass : Select main class of repetitive element
Subclass :
Gene Symbol :
Chromosome (D):


選擇特定種類的
重複性序列

GBRED 搜尋介面

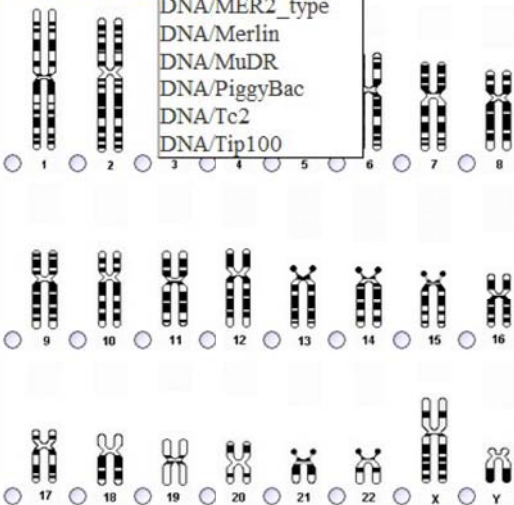
GBRED
[Homepage](#)
[Query](#)
■ [Manual](#)
■ [RepeatMasker](#)
CONTACT US
■ [E-mail](#)
VERSION
RepeatMasker version: 3.1.4
Rebase version : 20070801
Last updated : 20080928

QUERY September 30, 2008

Region :

RepeatClass :

Subclass : Select subclass of repetitive element which
Gene Symbol

Chromosome

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8
☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16
☐ 17 ☐ 18 ☐ 19 ☐ 20 ☐ 21 ☐ 22 ☐ X ☐ Y ☐ All

選擇細部種類的
重複性序列

GBRED 搜尋介面

GBRED

[Homepage](#)
[Query](#)
[Manual](#)
[RepeatMasker](#)

CONTACT US

[E-mail](#)

VERSION

RepeatMasker version: 3.1.4
Rebase version : 20070801
Last updated : 20080928

QUERY

Region :

All

RepeatClass :

All

Subclass :

All

Gene Symbol :

Enter the gene symbol to search relative repetitive element information. You can use the first few letters followed by * for a wildcard search. e.g.fgf*.

Chromosome (Default : All)

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

X

Y

All

Search

reset

輸入基因的名稱，
並可利用*當作萬
用字元來進行檢索

GBRED 搜尋介面

GBRED

[Homepage](#)
[Query](#)
■ [Manual](#)
■ [RepeatMasker](#)

CONTACT US

■ [E-mail](#)

VERSION

RepeatMasker version: 3.1.4
Repbase version : 20070801
Last updated : 20080928

QUERY

September 30, 2008

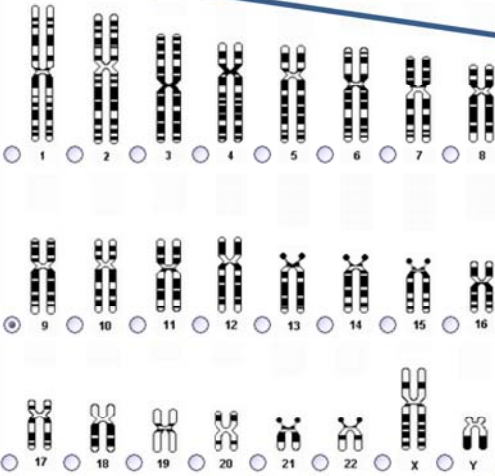
Region :

RepeatClass :

Subclass :

Gene Symbol :

Chromosome (Default : All)



1 2 3 4 5 6 7 8
9 10 11 12 13 14 15 16
17 18 19 20 21 22 X Y All

選擇要搜尋的染色體



UNLab
lar Genetics & Bioinformatics

GBRED 搜尋介面 (範例)

GBRED

Homepage
Query
■ Manual
■ RepeatMasker

CONTACT US

■ E-mail

VERSION

RepeatMasker version: 3.1.4
Repbse version : 20070801
Last updated : 20080928

QUERY September 30, 2008

Region :

RepeatClass :

Subclass :

Gene Symbol :

Chromosome (Default : All)

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20 ☐ 21 ☐ 22 ☐ X ☐ Y ☐ All

選擇3UTR

選擇LINEs

選擇LINE/CR1

以染色體9為例

UNLab
lar Genetics & Bioinformatics

GBRED 搜尋結果

前一頁所設定的搜尋條件

You have selected

- chromosome 09
- gene name is *
- region is 3UTR
- selected repeat is LINE/CR1

[Manual page](#)

基因正式縮寫

Search Result

			Total: 9
			Page: 1 / 1
Gene_ID	Gene_Symbol	Description	
ENSG0000023318	TXNDC4	Thioredoxin domain-containing protein 4 precursor (Endoplasmic reticulum resident protein ERp44). [Source:Uniprot/SWISSPROT;Acc:Q9BS26]	
ENSG00000095203	EPB41L4B	Band 4.1-like protein 4B (Protein EHM2) (FERM-containing protein CG1). [Source:Uniprot/SWISSPROT;Acc:Q9H329]	
ENSG00000137038	C9orf123	Uncharacterized protein C9orf123. [Source:Uniprot/SWISSPROT;Acc:Q96GE9]	
ENSG00000137073	UBAP2	Ubiquitin-associated protein 2. [Source:Uniprot/SWISSPROT;Acc:Q5T6F2]	
ENSG00000137135	C9orf100	VAV-like protein C9orf100. [Source:Uniprot/SWISSPROT;Acc:Q8N4T4]	
ENSG00000148053	NTRK2	BDNF/NT-3 growth factors receptor precursor (EC 2.7.10.1) (Neurotrophic tyrosine kinase receptor type 2) (TrkB tyrosine kinase) (GP145-TrkB) (Trk-B). [Source:Uniprot/SWISSPROT;Acc:Q16620]	
ENSG00000148153	C9orf80	Uncharacterized protein C9orf80. [Source:Uniprot/SWISSPROT;Acc:Q9NRY2]	
ENSG00000180471	C9orf31	Putative uncharacterized protein C9orf31 (Protein MOST-2). [Source:Uniprot/SWISSPROT;Acc:Q9NRJ2]	
ENSG00000203987		CDNA FLJ46557 fis, clone THYMU3039846. [Source:Uniprot/SPTREMBL;Acc:Q6ZR86]	

Page 1

Pages 1

(Total Data 9)

[New Query](#)

Ensembl所提供的
基因編號，點擊以
取得詳細資料

基因的全名及其來源

GBRED 搜尋結果

若搜尋條件中沒有選擇特定染色體，則
結果會列出各基因所在的染色體

You have selected

- gene name is *
- region is 3UTR
- selected repeat is LINE/CRT

[Manual page](#)

列出所在的染色體

連結到說明網頁

Search Result

[Next](#)

Total: 144

Page: 1 / 5

Chromosome	Gene_ID	Gene_Symbol	Description
1	ENSG00000006555	TTC22	Tetratricopeptide repeat protein 22 (TPR repeat protein 22). [Source:Uniprot/SWISSPROT;Acc:Q5TAA0]
1	ENSG00000009790	TRAF3IP3	TRAF3-interacting JNK-activating modulator (TRAF3-interacting protein 3). [Source:Uniprot/SWISSPROT;Acc:Q9Y228]
1	ENSG00000057252	SOAT1	Sterol O-acyltransferase 1 (EC 2.3.1.26) (Cholesterol acyltransferase 1) (Acyl coenzyme A:cholesterol acyltransferase 1) (ACAT-1). [Source:Uniprot/SWISSPROT;Acc:P35610]
1	ENSG00000077157	PPP1R12B	Protein phosphatase 1 regulatory subunit 12B (Myosin phosphatase- targeting subunit 2) (Myosin phosphatase target subunit 2). [Source:Uniprot/SWISSPROT;Acc:O60237]
1	ENSG00000097033	SH3GLB1	SH3 domain GRB2-like protein B1 (EC 2.3.1.-) (Endophilin-B1) (Bax- interacting factor 1) (Bif-1). [Source:Uniprot/SWISSPROT;Acc:Q9Y371]
1	ENSG00000117597	C1orf107	Digestive organ expansion factor homolog. [Source:Uniprot/SWISSPROT;Acc:Q68CQ4]
1	ENSG00000118473	SGIP1	SH3-containing GRB2-like protein 3-interacting protein 1 (Endophilin- 3-interacting protein). [Source:Uniprot/SWISSPROT;Acc:Q9BQI5]
1	ENSG00000133065	SLC41A1	solute carrier family 41 member 1 [Source:RefSeq_peptide;Acc:NP_776253]

Ensembl所提供的
基因編號，點擊以
取得詳細資料

在此以TTC22基因為例子，
點擊基因編號以取得基因的
詳細資料

基因詳細資訊

Ensembl基因編號，可連結至Ensembl網站

基因正式縮寫

基因的全名及其來源

Gene Content

Gene ID [ENSG00000006555](#)

Gene Symbol TTC22

Description Tetratricopeptide repeat protein 22 (TPR repeat protein 22). [Source:Uniprot/SWISSPROT;Acc:Q5TAA0]

Strand -1

Biotype protein_coding

Gene Location Chromosome 01 location 55019504-55039459

Transcript

Transcript ID	Transcript start	Transcript end
ENST00000371274	55022813	55039459

此基因位於染色體的正股或是負股

Please select the transcript ID for further information.

[New Query](#)

基因功能

基因在染色體上的位置

此基因所具有的 transcripts，點擊可顯示重複性序列的詳細資料

重複性序列詳細資訊

重複性序列所在的區域以及此區域長度

所選擇的transcript ID

Transcript	Transcript ID	Transcript start	Transcript end
	ENST00000371274	55022813	55039459

Graph Table - [ENST00000371274](#)

不同的色塊代表不同種類的重複性序列

Legend - Description of repetitive elements

SINEs	LTR elements	Small RNA	Simple repeats	Unclassified: Other	Transcription Start site
LINEs	DNA elements	Satellites	Low complexity	Unclassified: Unknown	Transcription End site

重複性序列的種類

3'UTR Length: 994bp

Graph

Query Region	Match Type	Repeat Name	Repeat Class	Hit Region
47..112	C	L3	LINE/CR1	4435..4368
588..713	C	L3	LINE/CR1	4347..4209

重複性序列在此區域的位置

重複性序列對應到保留序列的正股或負股

此段重複性序列的名稱

重複性序列對應到保留序列的位置

[New Query](#)

GBRED - 圖像列表

點擊圖中的重複性序列，可以fasta格式下載其序列

>ENST00000371274|3UTR|588..713

```
CTCACTGTCTCTTGATGCAGTC
CACTGCATAGTCATTCAGCTCT
GCAAGTTACACACTTAGGTGA
GTTCTGCCTCCTGCTAGCTTC
CCGCACTGGTCCCAGACCTGC
CTCCTAGAGCCCACACAGA
```

ENST00000371274

Description of repetitive elements

Classified: Other

Transcription
Start site

Classified: Unknow

Transcription
End site

Graph

Query Region

Match Type

Repeat Name

Repeat Class

Hit Region

47..112

C

L3

LINE/CR1

4435..4368

588..713

C

L3

LINE/CR1

4347..4209

藉由圖示呈現出
重複性序列分布
的相對位置

[New Query](#)

Visitor counts



GBRED

Gene-Based Repetitive Element Database

GBRED

[Homepage](#)

[Query](#)

■ [Manual](#)

■ [RepeatMasker](#)

CONTACT US

■ [E-mail](#)

VERSION

RepeatMasker version: 3.1.4

Rephase version: 20070801

Last updated: 20120905

Counter: 517

INTRODUCTION

GBRED (Gene-Based Repetitive Element Database) is a dedicated database containing repetitive element information which focus on the sequences flanking all human genes, such as 5' untranslated region (5'UTR), and 3' untranslated region (3'UTR). Recent articles imply that repetitive elements which located in these regions could regulate gene expression. However, there is no systematical survey to analyze which kind of repetitive elements prefer to locate in these regions and what is the role about these repetitive elements. GBRED is trying to provide this information to users. You can click the query button in left panel to enter the information query system.

NLab

Molecular Genetics & Bioinformatics

Thank you for your attention

