

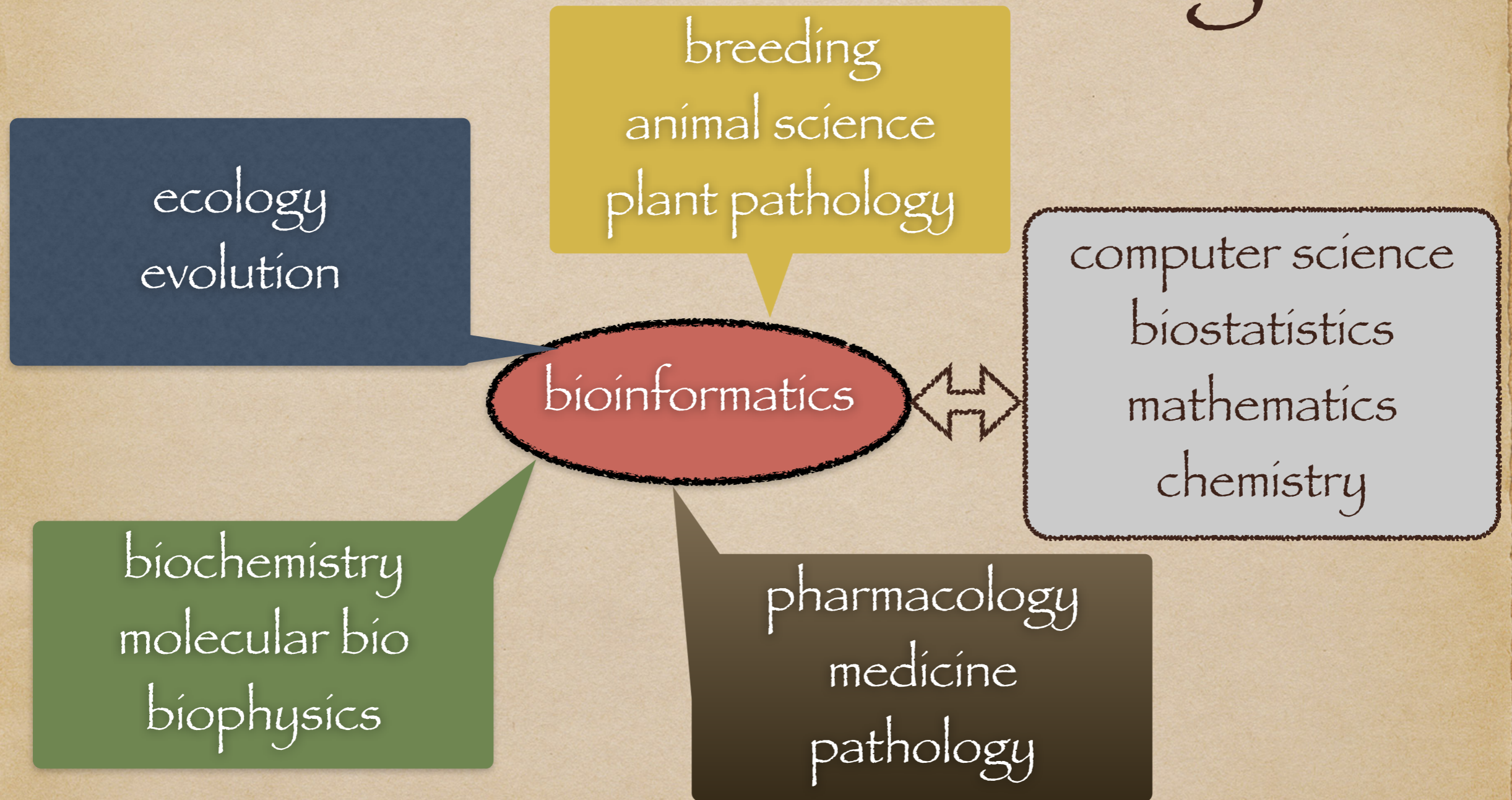
Thailand Bioinformatics: Research and Applications

Sissades Tongsimá

Bioinformatics Laboratory, BIOTEC

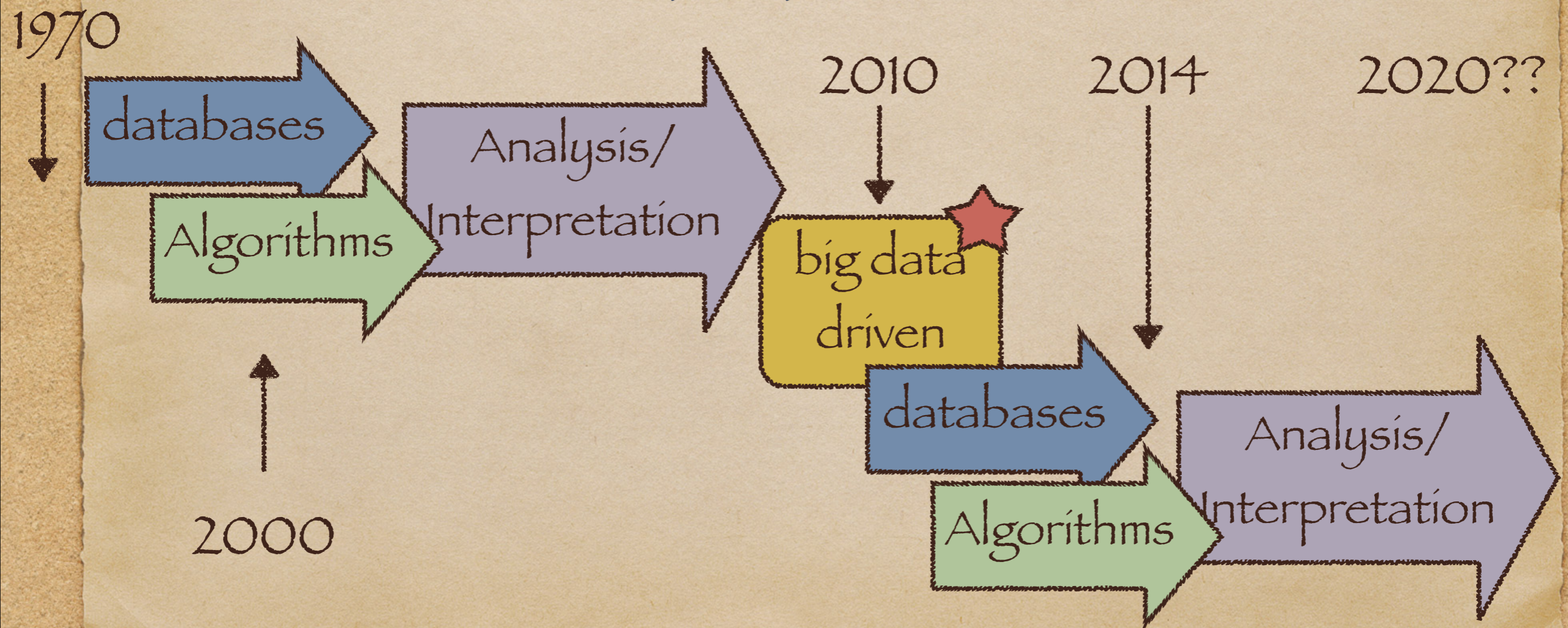
- ◆ Merging Computer Science, Biology, Information technology into one discipline ➔ Bioinformatics
- ◆ Not biological computation (to compute by imitating biological processes)

Bioinformatics Bridge



Commonly known bioinformatics

- ◆ Goal: “enable” the discovery of new insights by means of “global perspective



Bioinformatics is the future for bio-related disciplines

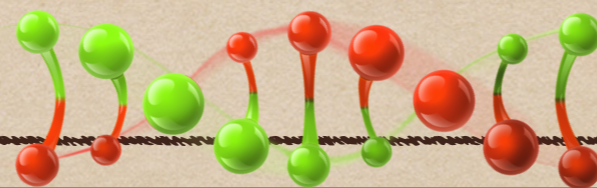
- ◆ Applications driven by next generation sequencing technology (gene mapping, pharmacogenomics etc.)
- ◆ Peptide inhibitor/drug/vaccine design
- ◆ Marker assisted selection (MAS) breeding

How to do "bioinformatics"

Biological "domain" experts



Bioinformatics/Computational biology



Biological drivers

Resources



Types of work (in Thailand)

- ◆ Data management or "Databases"

Bio-annotation
graphic visualization

- ◆ Novel algorithms for data manipulation

Easy tool
Pattern Recognition
Systems modeling

- ◆ Computational Biology

Applying tools and DB
New Computational workflow

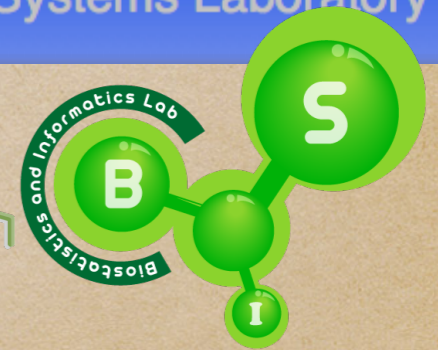
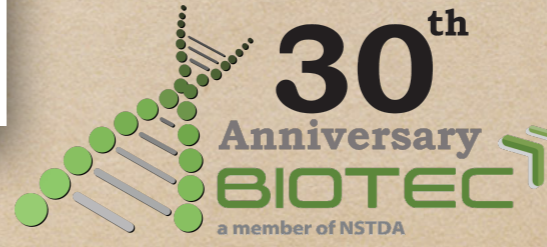
Master's Degree in
**Bioinformatics &
Systems Biology**



Thailand Bioinformatics



Information Systems Laboratory



Snapshots of Activities

- ◆ **KMUTT-CMU** (Bioinformatics program)
 - ◆ Starch biosynthesis (Prof. Supaporn)
 - ◆ Systems biology (Prof. Asawin/Jeerayuth)
- ◆ **Prince of Songkla University**
 - ◆ Shrimp immune/ Oil palm (Prof. Amornrat)
 - ◆ Bioinformatics (Prof. Unitsa/Pitipol)

Snapshots of Activities

- ◆ Chulalongkorn University
 - ◆ LINE-1/Cancer research (Prof. Apiwat)
 - ◆ Gene Mapping (Prof. Vorasuk)
 - ◆ Algorithm development (Prof. Chidchanok)
 - ◆ Molecular modeling/docking (Prof. Supoj)
 - ◆ Metagenomics (Prof. Naraporn)

Snapshots of Activities

- ◆ *Kasetsart University*

- ◆ Rice breeding program (Prof. Apichart)

- ◆ Molecular modeling/docking (Prof. Supa)

- ◆ *Khonkaen University*

- ◆ Animal breeding (Prof. Monchai)

- ◆ Pharmacogenomics (Prof. Wichitra)

Snapshots of Activities

- ◆ Mahidol University
 - ◆ Dengue (Prof. Preeda/Prapat)
 - ◆ Pop genetics (Prof. Patcharee/Bhoom)
 - ◆ Sequencing (Siriraj hospital)
 - ◆ Virus, GWAS (Prof. Wasun)
 - ◆ Thalassemia (Prof. Suthat)

What have we done?

Gene Information

Official symbol : [ITGAX](#) Full name : integrin, alpha X (complement component 3 receptor 4 subunit) Gene Id : 3687
 Gene type : protein-coding Position : Chr16 [31274010 - 31301819 bp] Size : 27810 bp Cytogenetic band : 16p11.2

Genomic regions, transcripts and products

Contig accession	mRNA accession	Protein accession	Exons	Orientation	Graphical View
NT_010393.15	NM_000887.3	NP_000878.2	30	forward	

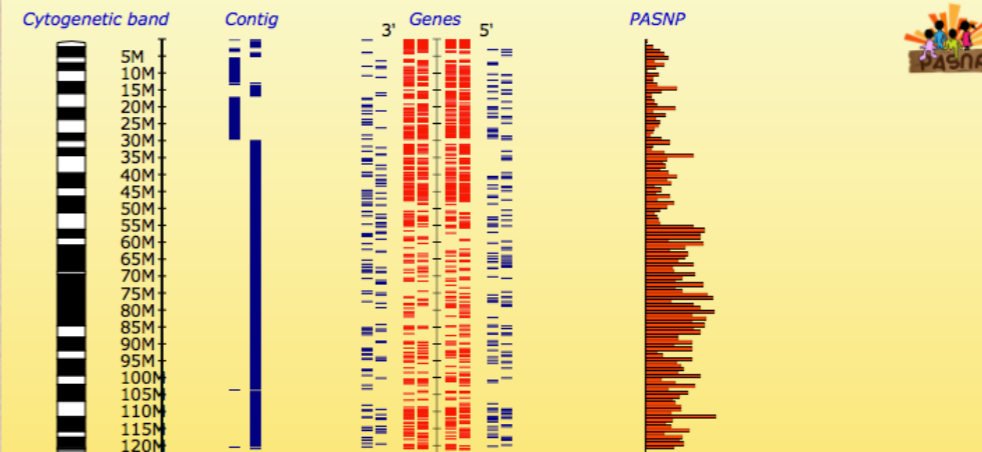
SNPs Information in region Chr16 : 31274010 - 31301819 (27810 bp)

Click the SNPs for detail information (Total 134 snps)

SNP ID	Observed	Validated	Chr Position	Strand	Fxn on Refseq
1. rs9938765	C/T	✓	Chr16:31274259	+	intron
2. rs2230424	C/T	✓	Chr16:31274819	+	missense
3. rs7190997	C/T	✓	Chr16:31275679	+	
4. rs59087566	A/T	✗	Chr16:31275780	+	
5. rs11574634	C/T	✓	Chr16:31275806	+	
6. rs61761294	A/G	✓	Chr16:31276006	+	
7. rs61761295	C/T	✓	Chr16:31276080	+	
8. rs56868377	C/G	✗	Chr16:31276305	+	
9. rs11574635	C/G	✓	Chr16:31276317	+	
10. rs56906552	C/T	✗	Chr16:31276324	+	
11. rs11574636	C/T	✓	Chr16:31276336	+	
12. rs11574637	C/T	✓	Chr16:31276375	+	
13. rs11150615	C/G	✓	Chr16:31277304	+	
14. rs11861211	G/T	✓	Chr16:31277395	+	
15. rs11150616	G/T	✓	Chr16:31277591	+	

dbSNP, HapMap, JSNP, PASNP, ThaiSNP1, ThaiSNP2, Thai depression-control, Thai drug related genes (TCELS)

Chromosome1



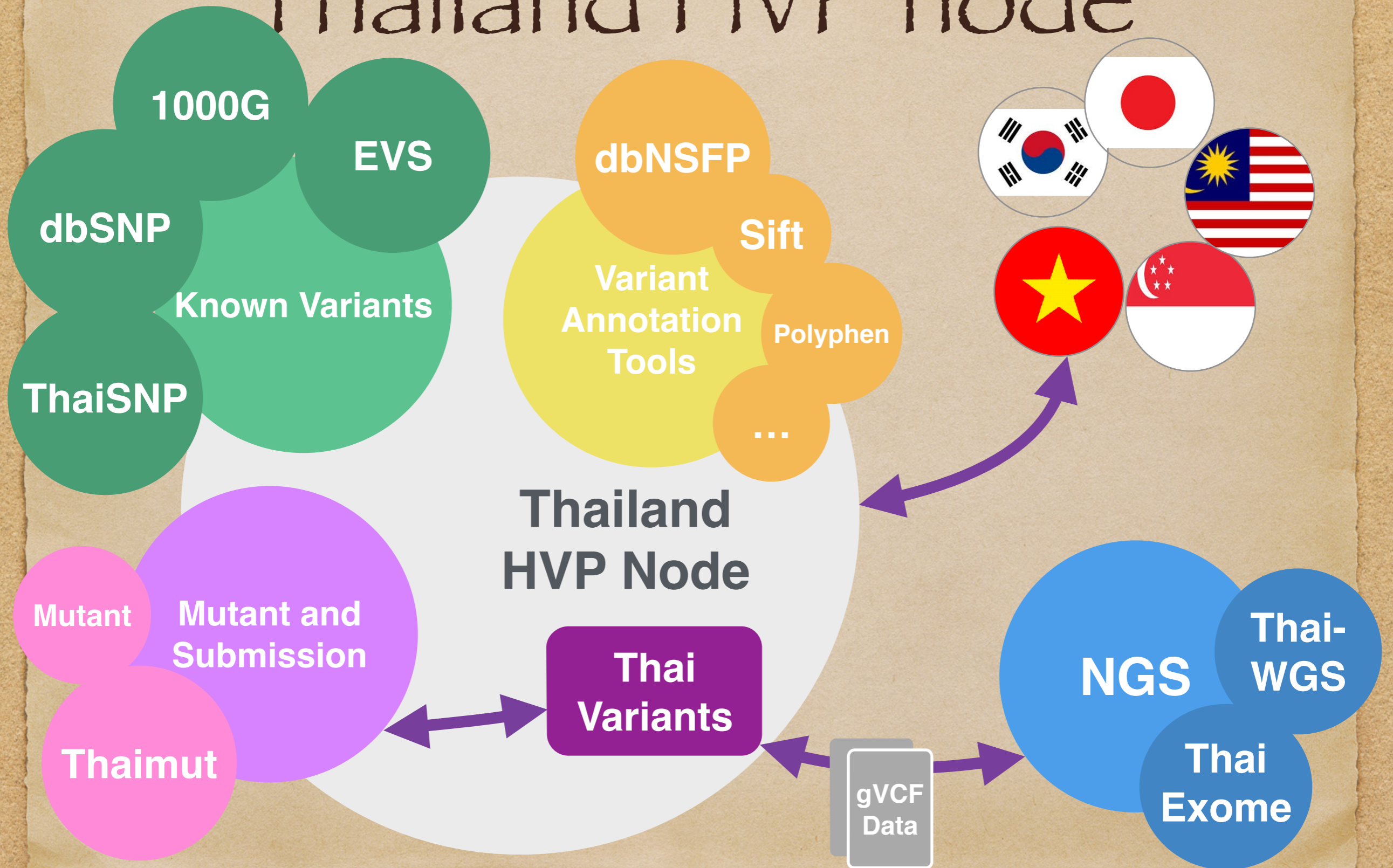
Database Summary

Thailand Human Mutation Database Information	Thailand Human Variation Database Information
all mutations : 589	all SNPs : 1536
mutated genes : 119	number of genes containing SNPs : 228
phenotypes(diseases) : 119	SNPs grouped by function :
publications : 219	locus-region : 133
mutations grouped by type :	coding-synon : 54
substitution : 507	coding-nonsynon : 39
deletion : 59	mRNA-utr : 203
insertion : 20	intron : 1135
indels : 3	near-gene : 45
	(Sum of these SNPs > all SNPs due to multiple gene isoforms.)

Data management

Genetic Variation databases

Thailand HVP node



MUTANT Mutation Annotation Tool

Gene HBA2 on chr.16
 NM_000517
 NP_000508
 NT_037887

NT_037887
 163500 162000 162000
 Chr 16 forward

substitutions

substitute position

base 1 T

base change to

Reset Save

Navigate to exon number

Genomic Ref. 715 720 725
 C A A R T A C C G T T A R G C T G

Coding Ref. 420 425

VarDetect 200601251500

SequenceViewer Window AnnotationViewer Window IndividualViewer Window TraceViewer Window AnnotationMapViewer Window

Remove Clear SNP

19010 SNP(A/G):96.87(3.12,sample=64,type=SNP
 19129 SNP(A/G):79.68(20.31,sample=64,type=SNP base ch.
 19248 SNP(G/G):70.31(29.68,sample=64,type=SNP
 19268 SNP(C/T):45.31(54.68,sample=64,type=SNP base ch.
 19627 SNP(C/G):96.87(3.12,sample=64,type=SNP
 19650 SNP(A/C):90.0(9.0,sample=60,type=SNP base change

Attributes Values

Offset 19368
 Ref-allele C
 Major/Minor 7/C
 Frequency 0.453125(0.546875
 Validated false
 Samples 64
 Type =SNP base change

WholeMapViewer Window

VarDetect NCS 201306052037

Variant Table (tab: default)

CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	DP	GI
chr1	788844	.	C	G	33	PASS	NS=275; CT=QDP; 0:1316; 0:1337			
chr1	788903	.	C	A	34	PASS	NS=275; CT=QDP; 0:1343; 0:1343			
chr1	788945	.	G	C	29	PASS	NS=275; CT=QDP; 0:1307; 0:1291			
chr1	788948	.	C	C	31	PASS	NS=275; CT=QDP; 0:1332; 0:1305			
chr1	788952	.	C	A	34	PASS	NS=275; CT=QDP; 0:1344; 0:1349			
chr1	788987	.	C	C	35	PASS	NS=275; CT=QDP; 0:1357; 0:1361			
chr1	789001	.	C	A	35	PASS	NS=275; CT=QDP; 0:1345; 0:1371			
chr1	789043	.	C	C	35	PASS	NS=275; CT=QDP; 0:1367; 0:1358; 8.4			
chr1	789050	.	C	A	33	PASS	NS=275; CT=QDP; 0:1358; 0:1348; 8.8			
chr1	789095	.	C	G	36	PASS	NS=275; CT=QDP; 0:1381; 0:1342			
chr1	789099	.	C	A	36	PASS	NS=275; CT=QDP; 0:1371; 0:1382			
chr1	789144	.	C	A	36	PASS	NS=275; CT=QDP; 0:1384; 0:1387			
chr1	789193	.	C	G	33	PASS	NS=275; CT=QDP; 0:1345; 0:1327			
chr1	803432	.	T	C	34	PASS	NS=275; CT=QDP; 0:1344; 0:1347			
chr1	803596	.	T	C	37	PASS	NS=275; CT=QDP; 0:1375; 0:1378			
chr1	803688	.	C	G	37	PASS	NS=275; CT=QDP; 0:1385; 0:1377			
chr1	803702	.	C	A	37	PASS	NS=275; CT=QDP; 0:1385; 0:1377			
chr1	804073	.	C	A	38	PASS	NS=275; CT=QDP; 0:1383; 0:1384			
chr1	804090	.	C	T	38	PASS	NS=275; CT=QDP; 0:1382; 0:1383			
chr1	804107	.	T	C	38	PASS	NS=275; CT=QDP; 0:1381; 0:1372			
chr1	804138	.	A	T	36	PASS	NS=275; CT=QDP; 0:1371; 0:1352			
chr1	804150	.	A	G	36	PASS	NS=275; CT=QDP; 0:1380; 0:1352			
chr1	804177	.	A	C	37	PASS	NS=275; CT=QDP; 0:1388; 0:1378			
chr1	804187	.	A	C	37	PASS	NS=275; CT=QDP; 0:1382; 0:1372			
chr1	804197	.	C	T	34	PASS	NS=175; CT=QDP; 1:1343; 0:1354			

2D Feature Matrix (default)

DP: 27 34 26 37 25 35 10 37 12 35 18 35 16 36 67 35

GI: 27 34 26 37 25 35 10 37 12 35 18 35 16 36 67 35

PropertyComponent Window

CHROM: chr1
 POS: 789256
 REF: T
 ALT: C
 QUAL: 35
 FILTER: PASS
 INFO: NS=275; 2-DP=103; AF=1.00; C
 FORMAT: CT; CQ; DPAA
 DP: 11; 36; 79; 79
 S1: 1; 1; 35; 128; 128



Gene Information : ADAM33 (GeneID : 80332)

Organism : Homo sapiens (Human)
 Contig accession : NT_011387.8 mRNA accession : NM_025220.2 Protein accession : NP_079496.1 Orientation : reversed Number of exon : 22
 Click to Show/Hide Gene/Pseudogene view

3' Reverse Organism Human | Gene ADAM33 on mRNA NM_025220.2 | prot NP_079496.1 | contig NT_011387.8 | Chr 20 (22 exons, 14119 bp) 5'

3596200 3596200 3596500 3598800 3599300 3599400 3599700 3600000 3602300 3602300 3612500

Chr 20

locus region intron mRNA/Coding exon UTR exon

Select region on gene to design

1. Design primer for whole gene

2. Design primer for some region (By input Intron/Exon range number or user input region)

3. Design primer for whole gene amplification (primer must be designed in these specific region)

Whole Gene Exons only

Whole region Exons only From exon : 1 to exon : 22

List of regions (Example)

Left primer region range _____ bp

Right primer region range _____ bp

WASP Web-based Allele Specific Primer designing tool

Home | About Us | Contact | Help

Reminder: WasP tool requires a SVG-enable browser to view its content. If you are not using either Firefox or Opera to view this page, please download and install SVG plug-in here. Note: Microsoft IE has been reported as not compatible with viewing this web-site.

Chromosome Classify

Metaphase : 2_02-06_070113142503

Individual Chromosome : 31

Class - 1 (Straight individual chromosome): 29

Class - 2 (Bending individual chromosome): 2

Class - 3 (Overlapping chromosomes): 5

Class - 4 (Artifacts): 2

Total Segment : 38

Section A: Allele-Demo1 Demo2

Search by SNP ID

(Continue Section C if you

Tools Development

Standalone/Web tools



Methodology article

Iterative pruning PCA improves resolution of highly structured populations

Apichart Intarapanich¹, Philip J Shaw², Anunchai Assawamakin^{3,2}, Pongsakorn Wangkumhang², Chumpol Ngamphiw², Kridsakorn Chaichoompu², Jittima Piriyaopngsa² and Sissades Tongshima^{*2}

Address: ¹NECTEC 112 Thailand Science Park, Paholyothin Road, Klong 1, Klong Luang, Pathumtani 12120, Thailand, ²BIOTEC 113 Thailand Science Park, Paholyothin Road, Klong 1, Klong Luang, Pathumtani 12120, Thailand and ³Division of Molecular Genetics, Department of Research and Development, Faculty of Medicine Siriraj Hospital, Bangkok 10700, Thailand

Email: Apichart Intarapanich - apichart.int@nectec.or.th; Philip J Shaw - philip@biotec.or.th; Anunchai Assawamakin - anunchai_ice@yahoo.com; Pongsakorn Wangkumhang - pongsakorn.wan@biotec.or.th; Chumpol Ngamphiw - chumpol.nga@biotec.or.th; Kridsakorn Chaichoompu - kridsakorn.cha@biotec.or.th; Jittima Piriyaopngsa - jittima.pir@biotec.or.th; Sissades Tongshima* - sissades@biotec.or.th

* Corresponding author

Research Article

Biomarker Selection and Classification of "-Omics" Data Using a Two-Step Bayes Classification Framework

Anunchai Assawamakin,¹ Supakit Prueksaaron,² Supasak Kulawonganuchai,³ Philip James Shaw,³ Vara Varavithya,⁴ Taneth Ruangrajitpakorn,⁵ and Sissades Tongshima³

¹ Department of Pharmacology, Faculty of Pharmacy, Mahidol University, 447 Sri-Ayuthaya Road, Rajathevi, Bangkok 10400, Thailand

² Department of Electrical and Computer Engineering, Faculty of Engineering, Thammasat University, 99 Phahonyothin Road, Khlong Nueng, Khlong Luang, Pathum Thani 12120, Thailand

³ National Center for Genetic Engineering and Biotechnology, 113 Thailand Science Park, Phahonyothin Road, Khlong Nueng, Khlong Luang, Pathum Thani 12120, Thailand

⁴ Department of Electrical and Computer Engineering, King Mongkut University of Technology North Bangkok, 1518 Piboonsongkarn Road, Bangkok 10800, Thailand

⁵ Language and Semantic Technology Laboratory, National Electronic and Computer Technology Center, 112 Thailand Science Park, Phahonyothin Road, Khlong Nueng, Khlong Luang, Pathum Thani 12120, Thailand

Correspondence should be addressed to Sissades Tongshima: sissades@biotec.or.th

Limpiti et al. BMC Bioinformatics 2011, 12:255
http://www.biomedcentral.com/1471-2105/12/255

Algorithm Development ORIGINAL ARTICLE

METHODOLOGY ARTICLE

Open Access

Study of large and highly stratified population datasets by combining iterative pruning principal component analysis and structure

Tulaya Limpiti¹, Apichart Intarapanich², Anunchai Assawamakin³, Philip J Shaw³, Pongsakorn Wangkumhang², Jittima Piriyaopngsa³, Chumpol Ngamphiw³ and Sissades Tongshima^{3*}

Extracting predictive SNPs in Crohn's disease using a vacillating genetic algorithm and a neural classifier in case-control association studies

Khantharat Anekboon^a, Chidchanok Lursinsap^{a,*}, Suphakant Phimoltares^a, Suthat Fucharoen^b, Sissades Tongshima^c

^a Advanced Virtual and Intelligent Computing (AVIC) Research Center Department of Mathematics and Computer Science, Chulalongkorn University Bangkok, Thailand

^b Institute of Science and Technology for Research and Development Mahidol University, Nakhonpathom, Thailand

^c Genome Institute, National Center for Genetic Engineering and Biotechnology, Pathumtani, Thailand

iLOCi: a SNP interaction prioritization technique for detecting epistasis in genome-wide association studies

Jittima Piriyaopngsa¹, Chumpol Ngamphiw¹, Apichart Intarapanich², Supasak Kulawonganuchai¹, Anunchai Assawamakin¹, Chaiwat Bootchai¹, Philip J Shaw¹, Sissades Tongshima^{1*}

From Asia Pacific Bioinformatics Network (APBioNet) Eleventh International Conference on Bioinformatics InCoB2012

Bangkok, Thailand. 3-5 October 2012

Abstract

Background: Genome-wide association studies (GWAS) do not provide a full account of the heritability of genetic diseases since gene-gene interactions, also known as epistasis are not considered in single locus GWAS. To address this problem, a considerable number of methods have been developed for identifying disease-associated gene-gene interactions. However, these methods typically fail to identify interacting markers explaining more of the disease heritability over single locus GWAS, since many of the interactions significant for disease are obscured by uninformative marker interactions e.g., linkage disequilibrium (LD).

Results: In this study, we present a novel SNP interaction prioritization algorithm, named iLOCi (Interacting Loci).

Journal of Human Genetics (2009) 54, 634-641
© 2009 The Japan Society of Human Genetics All rights reserved 1434-5161/09 \$32.00
www.nature.com/jhg

ORIGINAL ARTICLE

pHCR: a Parallel Haplotype Configuration Reduction algorithm for haplotype interaction analysis

Wattanan Makarasara^{1,2,3,4}, Natsuhiko Kumasaka³, Anunchai Assawamakin^{4,5}, Atsushi Takahashi³, Apichart Intarapanich⁶, Chumpol Ngamphiw⁴, Supasak Kulawonganuchai⁴, Uttapong Ruangrit⁴, Suthat Fucharoen^{1,2}, Naoyuki Kamatani³ and Sissades Tongshima⁴

Insight into the Peopling of Mainland Southeast Asia from Thai Population Genetic Structure

Pongsakorn Wangkumhang^{1,2*}, Philip James Shaw^{1*}, Kridsakorn Chaichoompu¹, Chumpol Ngamphiw^{1,2}, Anunchai Assawamakin³, Manit Nuinoon⁴, Orapan Sripichai⁵, Saovaros Svasti⁵, Suthat Fucharoen⁵, Verayuth Praphanphoj⁶, Sissades Tongshima^{1*}

J Mol Model

DOI 10.1007/s00894-012-1565-2

ORIGINAL PAPER

Molecular dynamics simulations reveal structural instability of human trypsin inhibitor upon D50E and Y54H mutations

Wanwimon Mokmak · Surasak Chunsrivirok · Anunchai Assawamakin · Kiattawee Choowongkomon · Sissades Tongshima

BMC Genomics

Proceedings

Prediction of avian influenza A binding preference to human receptor using conformational analysis of receptor bound to hemagglutinin

Nipa Jongkon¹, Wanwimon Mokmak², Daungmanee Chuakheaw¹, Philip J Shaw², Sissades Tongshima^{*2} and Chak Sangma^{*1}

PROCEEDINGS

Open Access

Metagenomic profiles of free-living archaea, bacteria and small eukaryotes in coastal areas of Sichang island, Thailand

Naraporn Somboonna^{1*}, Anunchai Assawamakin², Ali Sissades Tongshima²



Mapping Human Genetic Diversity in Asia
The HUGO Pan-Asian SNP Consortium
Science **326**, 1541 (2009);
DOI: 10.1126/science.1177074

Sissades Tongshima

2012; doi:10.1038/jhg.2012.52

BMC Genetics

This Provisional PDF corresponds to the article as it appeared upon acceptance. Fully formatted PDF and full text (HTML) versions will be made available soon.

Population structure of four Thai Indigenous chicken breeds

BMC Genetics 2014, 15:40 doi:10.1186/1471-2156-15-40

Supamit Mekchay (agismkch@chiangmai.ac.th)
Pantaporn Supakankul (pantaporn@gmail.com)
Anunchai Assawamakin (anunchaiice@gmail.com)
Alisa Wilantho (alisa.wil@biotec.or.th)
Wanwisa Chareanchim (wanwisa.klu@biotec.or.th)
Sissades Tongshima (sissades@biotec.or.th)

OPEN ACCESS Freely available online

PLOS ONE

Hypomethylation of Intragenic LINE-1 Represses Transcription in Cancer Cells through AGO2

Chatchawit Apornthewan^{1*}, Chureerat Phokaew^{2*}, Jittima Piriyaopongsa^{3*}, Chumpol Ngamphiw^{3*}, Chupong Ittiwut⁴, Sissades Tongshima³, Apiwat Mutirangura^{4*}

¹ Department of Mathematics, Faculty of Science, Chulalongkorn University, Bangkok, Thailand, ² Inter-Department Program of BioMedical Sciences, Faculty of Graduate School, Chulalongkorn University, Bangkok, Thailand, ³ National Center for Genetic Engineering and Biotechnology, Genome Institute, Thailand Science Park, Pathumtani, Thailand, ⁴ Department of Anatomy, Faculty of Medicine, Center of Excellence in Molecular Genetics of Cancer and Human Diseases, Chulalongkorn University, Bangkok, Thailand



OPEN ACCESS Freely available online

PLOS ONE

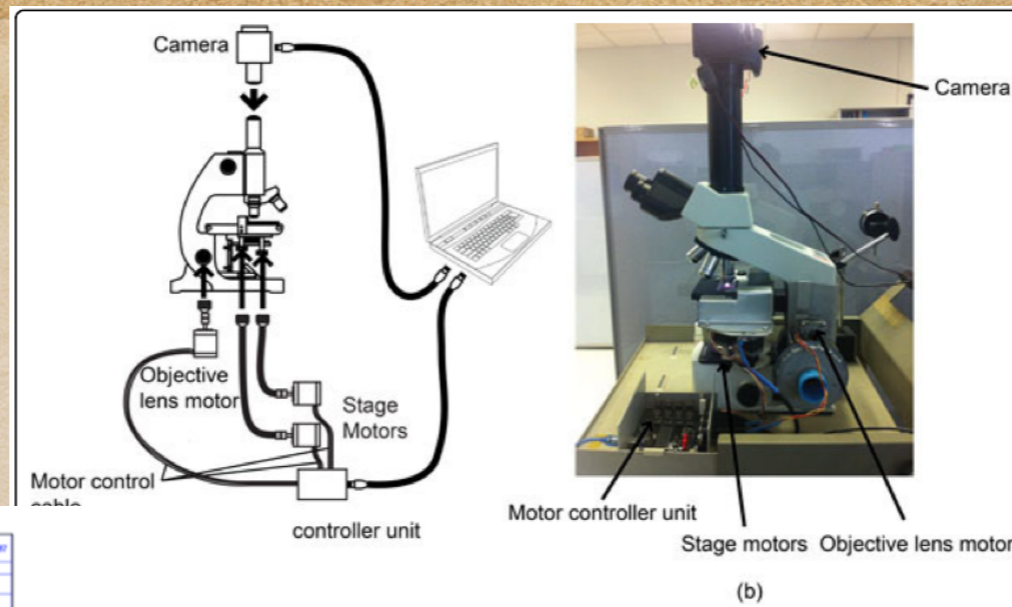
Open Access

Distal Effect of Amino Acid Substitutions in CYP2C9 Polymorphic Variants Causes Differences in Interatomic Interactions against (S)-Warfarin

Panida Lertkiatmongkol^{1,2,3}, Anunchai Assawamakin⁴, George White³, Gaurav Chopra^{3,5}, Pornpimol Rongnoparut¹, Ram Samudrala³, Sissades Tongshima^{2*}

COMMENTARY

An abundance of population-specific monomorphic SNPs may or may not be meaningful: a commentary on differences in allele frequencies of familial hypercholesterolemia SNPs in the Malaysian



Motor control cable controller unit
Motor controller unit Stage motors Objective lens motor
(b)

Motorized units for objective lens and stage movement. (a) The schematic of automate microscope system, a motor controller unit and an eyepiece camera. There are three motorized units. Two of them control the stage and the objective lens motorized unit controls the objective lens position. An eyepiece camera is used for image analysis module in a computer via USB connection. (b) Photo of assembly of all units.

Uttamatin et al. *BMC Bioinformatics* 2013, **14**(Suppl 16):S13
<http://www.biomedcentral.com/1471-2105-14/S16/S13>

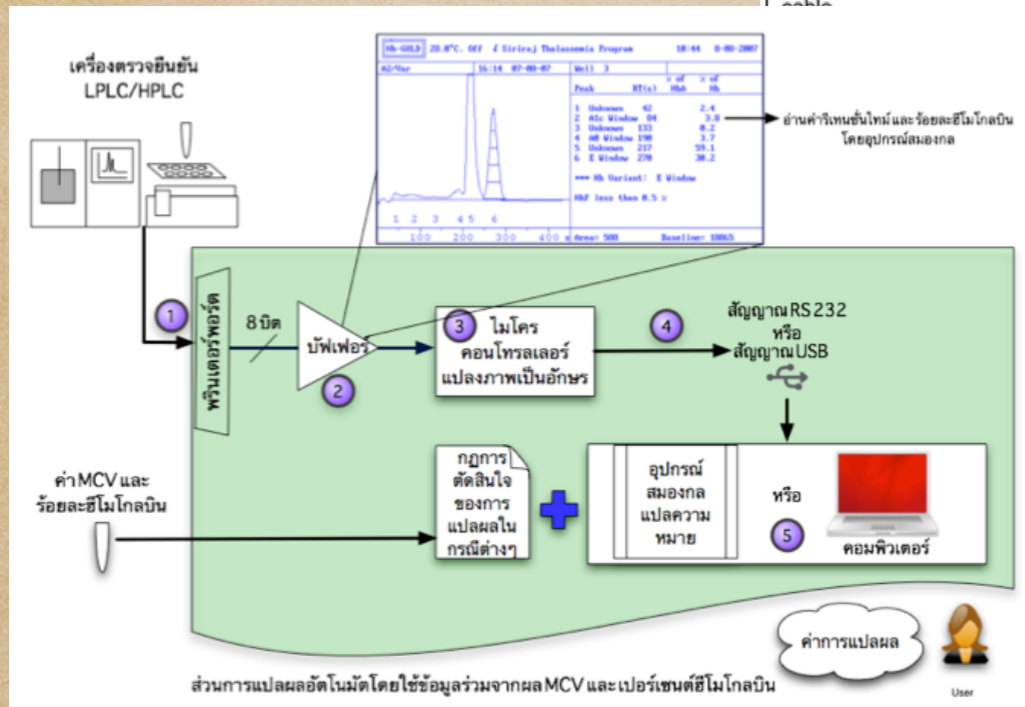


RESEARCH

Open Access

MetaSel: a metaphase selection tool using a Gaussian-based classification technique

Ravi Uttamatin¹, Peerapol Yuvapoositanon², Apichart Intarapanich³, Saowaluck Kaewkamnerd³, Ratsapan Phuksaritanon⁴, Anunchai Assawamakin⁵, Sissades Tongsim^{5*}



Medical Informatics

Hb Typing interpretation/Malaria Image Proc/
 Karyotyping Image Proc

Summary

- ◆ Bioinformatics “**assists**” discovery of new insights
- ◆ Bioinformatics research (all three aspects) is **driven by the data**
- ◆ To be successful, the key components must be in place (no guarantee, though :-)
- ◆ In Thailand, NGS could be a turned-key technology

Acknowledgement

- ◆ Members from Bioinformatics Lab, BIOTEC
- ◆ Prof. Suthat Fucharoen (Mahidol)
- ◆ Prof. Apiwat Mutirangura (Chulalongkorn)
- ◆ Prof. Vorasuk Shotilersuk (Chulalongkorn)
- ◆ Prof. Chanin Limwongse (Siriraj Hospital)
- ◆ Every Thai bioinformaticians