Journal of Theoretical and Applied Information Technology
31st March 2017. Vol.95. No 6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

THE FIRST MODULAR DATABASE OF INDONESIAN GENES-ASSOCIATED DISEASES INFORMATION

[1,3,4] RIKA YULIWULANDARI, [2,3]NOVA EKA DIANA, [3]KINASIH PRAYUNI

[1]Department of Pharmacology, Faculty of medicine, YARSI University, Jakarta, Indonesia

[2]Department of Informatics, Faculty of Information Technology, YARSI University, Jakarta, Indonesia

[3]Genomic Medicine Research Group, YARSI Research Institute, YARSI University, Jakarta, Indonesia

[4] The Indonesian Pharmacogenomics Working Group

E-mail: rika.yuliwulandari@yarsi.ac.id, nova.diana@yarsi.ac.id

ABSTRACT

This paper aims to develop a modular schema database for gene-associated diseases incidence in Indonesia, which is called as Genomic Medicine Research Group Database (GMRGDB). A modular architecture is beneficial to facilitate any manipulation and addition of new genes-affiliated diseases in the future. Currently, GMRGDB contains two categories, N-acetyltransferase 2 (NAT2) and Human Leukocyte Antigen (HLA) databases. NAT2 database is a repository of NAT2 genes and anti-TB drugs association, while HLA database records diseases related to Human Leukocyte Antigen. GMRGDB consists of patients' information from various ethnics in Indonesia, such as Javanese, Melayu, Betawi, Sundanese, and Batak. Users can view a number of samples for each database, detailed representation of Single Nucleotide Polymorphism (SNP), and also filter the data for specified ethnic through GMRGDB browser.

Keywords: *Genes-Associated Disease, Modular Database, N-Acetyltransferase 2, Human Leukocyte Antigen, Indonesian Ethnics*

1. INTRODUCTION

Several genetic mapping studies of some ethnics in Indonesia, especially Javanese and Sundanese population, have been conducted such as Human Leucocyte Antigen (HLA) and N-Acetyltransferase 2 (NAT2) studies [1, 2]. HLA is essential in identifying histocompatibility and disease susceptibility [3, 4]. The identification of HLA is also important for anthropological studies to elucidate genetic relationship among various ethnic groups in Indonesia [5]. Meanwhile, NAT2 has a major part to activate and deactivate arylamine and hydrazine drugs and carcinogen [6]. Polymorphisms in these genes are responsible for classifying the phenotype of NAT2 into a rapid, intermediate, and slow acetylator [6, 7].

The previous study of HLA in Javanese and Sundanese population showed the allele frequency of HLA-A, HLA-B, and HLA–DRB1 alleles; and also the broad relationship between these populations with other Asian inhabitants. Yuliwulandari et al reported that HLA allele types have an association with many diseases [1]. Previously, Yuliwulandari et al in their research also revealed that the slow acetylator of NAT2 variants were common in Indonesian populations, specifically within Javanese and Sundanese ethnics [2]. And in 2016, Yuliwulandari et al showed the association among those slow acetylators and susceptibility to Anti-Tuberculosis Drug Induced Liver Injury (AT-DILI) for Indonesian population [8].

There is yet a centralized database focusing on managing gene-associated diseases in Indonesia. Most of the data are still stationed in the disparate locations and there has not been an interconnected infrastructure to communicate the data. Hence, it is arduous for researchers to learn, analyze and find out the association of genes with the diseases. A centralized and comprehensive database for these data will definitely beneficial to researchers for cultivating more salient information and utilizing them on finding the solution of gene-associated diseases.

This paper aims to develop a modular database for representing gene-associated diseases in Indonesia. The organization of the article follows.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Section 1 describes the background of the research. Section 2 explains related works about methods that are used to develop this research. Next, Section 3 defines the methodologies for implementing the system. Section 4 contains the experiment results and its evaluations. Finally, Section 5 presents the conclusion and future works of our research.

2. RELATED WORKS

Many researchers in the world have adopting database techniques in information technology area to administer genes data. The National Center for Biotechnology Information (NCBI) is the most promising research institute on genes database area. This institute had conducted their researches since 2001 by creating dbSNP, a centralized database that integrated genome variations from various sources, such as GenBank, Pubmed, LocusLink and the Human Genome Project data [9]. In 2005, NCBI upgraded their LocusLink by adding Entrez Gene database which focused on the genomes that had been fully sequenced or were scheduled for intense sequence analysis. This database represented more detailed taxonomies of integrated data from NCBI's Reference Sequence (RefSeq) and other databases available in NCBI website [10]. NCBI kept working on updating and completing their database by appending various resources such Entrez, BLAST, RefSeq, UniGene, HomoloGene, the Conserved Domain Database (CDD), Protein Clusters, etc. Detail explanation and references about their services were described in the NCBI documentation that was released on 2013 and 2015 [11, 12].

Japan, Singapore, and Malaysia are three of the most active countries that conducting genomes-database research in Asia. Some researchers in Japan Science and Technology (JST) and The University of Tokyo had started to collect Japanese genomes variation since 2000. They built JSNP, a repository of 150,000 Japanese Single Nucleotide Polymorphism (SNP) that aim was to identify the association between the polymorphism and common disease of the drug reactions [13]. Further, many institutions collaborated to develop more comprehensive Japanese genetic database.

Five institutions: Kyoto University, Tohoku University, National Research Institute for Child Health and Development, the University of Tokyo, and Yokohama University, worked together to develop The Human Genetic Variation Database (HGVD). This database stored Japanese genetic variation of 1208 individuals. Browser interface of HGVD provides many utilities for viewing allele and genotype frequencies, eQTL significance, and also the number of samples [14]. Another research, The Malaysian Node of the Human Variome Project (MyHVP) had started on collecting genotype and phenotype variation of Malaysian population since 2010. This project focused on managing ethnic-specific database; hence the clinicians and researchers can obtain up-to-date and accurate information on population-specific culture [15].

Complete and rigorous genomic database need a thorough process to identify entities and relationship among genes-variation and diseases. The ontology-based representation is beneficial to fully describe the association among entities. This approach had been widely used to portray much medical and biological information. Lawrence Berkeley National Laboratory and Harvard University created Chado to represent the ontology-based relational database schema of genome-associated biological information [16]. Ontology was also utilized to identify and represent the incidence and spread of Malaria disease in Indonesia. A thorough process of attributes and relationship identification was crucial to extract important information from a huge number of data (malaria data warehouse) [17, 18]. Besides an impeccable entities and relationship representation, a well-developed database should provide the capability to adapt with the amendment and the transformation of the increasing genomes data. Modular-database architecture is perfect to facilitate any modification and data manipulation in the future [16].

3. MATERIAL AND METHODS

2.1 Data Source

This project aims to create a dynamic database for representing the gene-associated diseases information in Indonesia comprehensively, titled "Genomic Medicine Research Group Database (GMRGDB)". Each disease has a number of features determine the condition of a patient, and it may share some common parameters or information with other diseases. The number of parameters and diseases might also increase in the future. Hence, the database should have the modularity features to accommodate any improvements and modifications.

In this research, currently we categorized the data into two categories based on the project type:

a. NAT2 Project contains the interconnection between Anti-Tuberculosis Drug-Induced

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Liver Injury (AT-DILI) enzyme called N-acetyltransferase 2 (NAT2) and anti-TB drugs.

b. HLA Project consists of diseases data which are associated with Human Leukocyte Antigen (HLA).

We have successfully collected a total of 495 gene-associated data of various ethnics in Indonesia from 2015-2016. These data were classified into a control and patient data. A control data shows no negative effect of the designated drugs to the disease. While a patient data reveals information about the negative association of the designated drugs to the disease. Table I summarizes our research data. NAT2 database contains a total of 50 controls and 173 patients information for Javanese, Sundanese, Betawi, Melayu and Batak ethnics. Meanwhile, HLA project consists of a total of 99 and 173 data for controls and patient data, respectively.

2.2 Data Features Identification

As aforementioned in Table I, there are two groups of gene-associated diseases in our dataset. Those data were classified as control and patient data. Each group has a number of features determine the patient condition. Some features may be similar for different groups. Fig. 1 illustrates the identified parameters for NAT2 and HLA projects. Entities name with a Control-prefix depicts the control data for each project, while a Patient-prefix denotes a patient data. Patient NAT2 and Patient HLA data share the same SNP attributes: sgot0, sgpt0, bilirubin0, sgot1, sgpt1, bilirubin1, bilirubinDt, history, familyHistory, and therapy parameter.

Table 1: Genes-Associated Diseases Data In Indonesia (2015-2016)

Project	Data Type		Ethnics
	Control	*Patient*	
NAT2	50	173	Javanese, Sundanese, Betawi, Melayu, Batak
HLA	99	173	Betawi, Javanese, Sundanese, Batak, Melayu

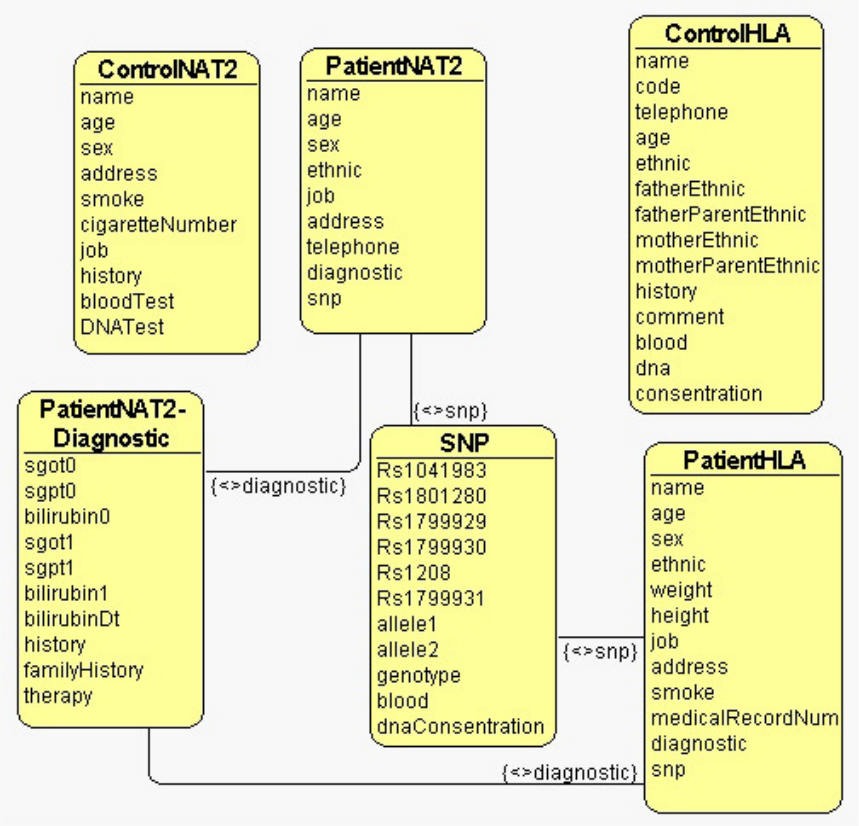


Figure 1: Project Features Identification

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Fig. 2 depicts the ontology of our current dataset that illustrates the interrelationships of the entities. The '*n*' symbol in the figure describes the common features owned by two or more entities. For example, 'blood' and 'dna' attributes are associated with ControlHLA and PatientHLA entity. Another parameters, 'history', 'familyHistory', and 'therapy', are also shared by PatientNAT2 and PatientHLA entity. There are many 'n' symbols in the figure, so it means that the entities have many similar features for identifying the designated disease. In this figure, there are some entities that come together to create a virtual group, and there is one entity that standing alone and does not associate with other entity (ControlNAT2 entity). Smaller the distance between two entities, then those entities are closer and share more features compare to other entity. In contrary, the apparent non-zero direct straight line connection of one entity to another, then this entity has unique parameters that do not belong to others.

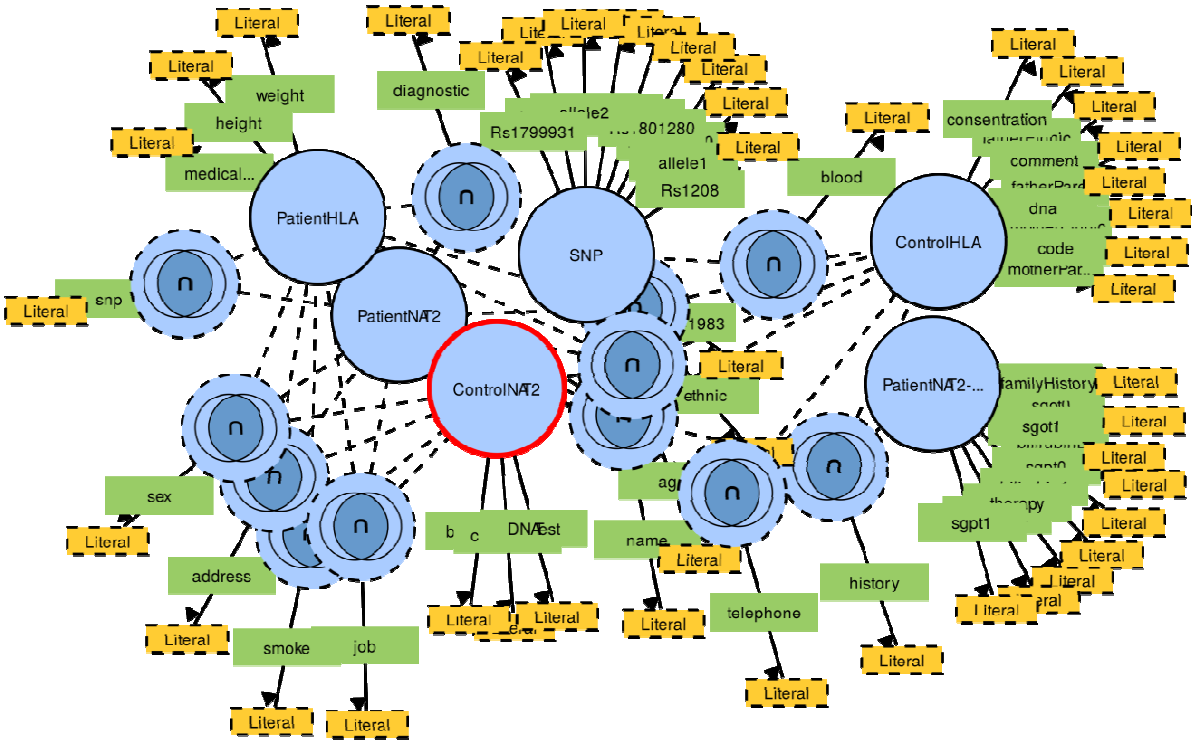


Figure 2: Ontology Of Genes-Associated Diseases

2.3 Database Construction and Implementation

In this research, some entities that represent gene-associated diseases have some similar features to diagnose the patient. Furthermore, there are some additional requirements to manage the gene-associated projects in our laboratory. Those requirements are:

a. The number of gene-associated projects may increase in the future

b. A project may have zero or more than one subproject

c. Each project and subproject have a set of attributes to identify the gene-associated disease

According to the aforementioned requirements, the proposed database should able to dynamically add a new project or subproject with a set of attributes. Fig. 3 illustrates the generic schema of our genes-associated disease projects. The main entities are PROJECTS, SUBPROJECTS, and USER. PROJECTS entity records a number of projects that are currently managed in our laboratory. PROJECT_DETAIL table saves the detail information of each project in PROJECTS table. A unique id and project name describe the project. Numerous attributes give a complete description of a project, which may be different for each. A project may have zero or more than one subprojects, in which each subproject have a collection of parameters that differ to their parent (a

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

project). List of all subprojects is summarized in SUBPROJECTS table. Each project is managed by a user that has been appointed as a person-in-charge (PIC) for this project. USER table stores more detail information for all users, including their role in using the system.

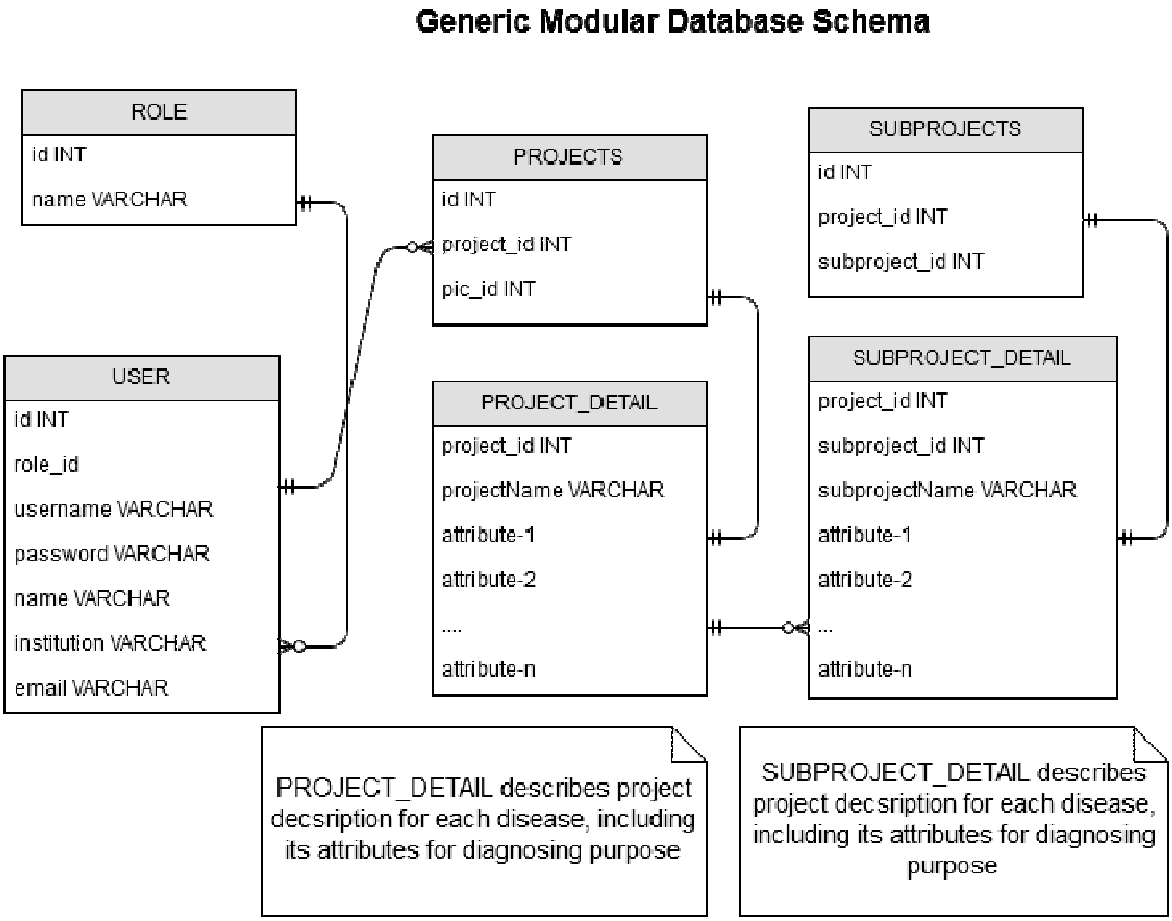


Figure 3: Generic Modular Schema Of Main Database

Fig. 4 shows the tree-tier architecture of our genes-associated disease database, which consists of the client, server, and database. We developed a web-client interface using Bootstrap framework which based on Hypertext Markup Language (HTML), Cascading Style Sheet (CSS), JavaScript, and Hypertext Processor (PHP) language programming. Apache Server was used to mediate the Structured Query Language (SQL) request from the client to our modular database, which was developed using MySQL.

4. RESULTS AND DISCUSSION

GMRGDB can be accessed by three types of user: database administrator, project-pic, and user portal. The project-pic is a user that the database administrator has assigned as a person-in-charge for managing the project. Once a user has been admitted as a project-pic, next he can add or manipulate the genes-associated diseases data. Those data can be accessed by all user portals, anyone who know the web application address and has an internet connection to access it.

The database is organized into two main categories:

a) **NAT2 database** contains patients' information of NAT2 enzyme and anti-TB drug association. These data was divided into a control and patient categories, which consist of Java (N=65), Betawi (N=27), Sundanese (N=12), Melayu (N=4), Batak (N=9), and unknown (N=55) ethnics.

b) **HLA database** stores a total of 272 patients data related to HLA-disease information, with Melayu (N=133), Batak (N=10), Limo Panjang (N=2), and unknown (N=48) ethnics.

Journal of Theoretical and Applied Information Technology
31st March 2017. Vol.95. No 6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

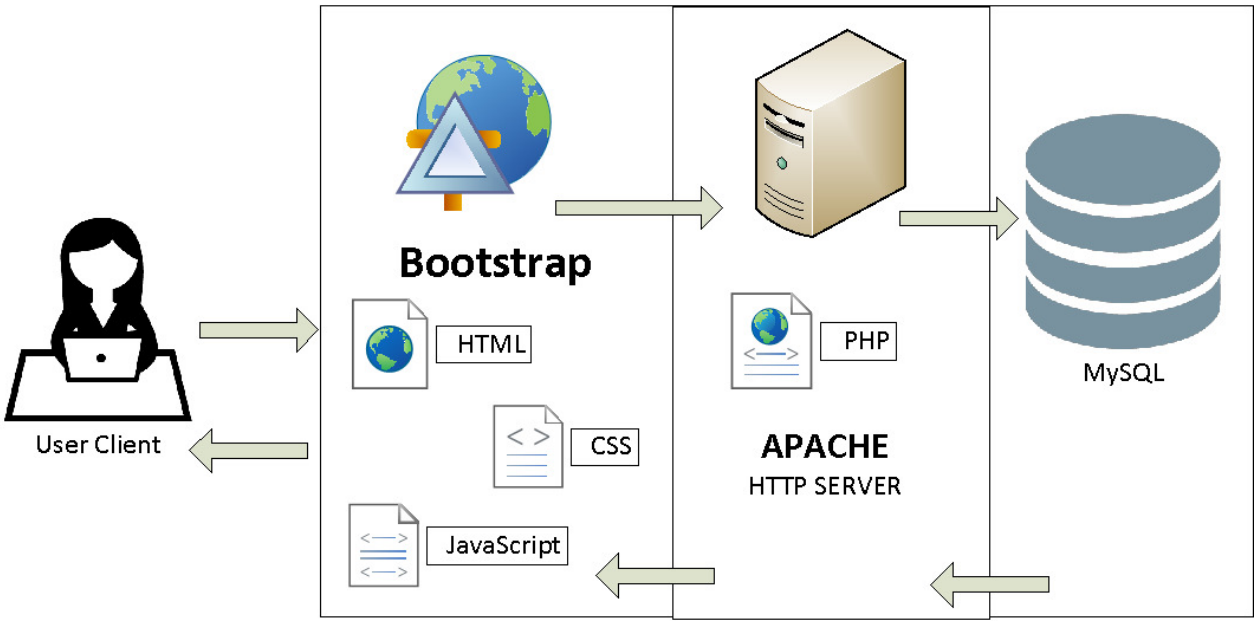
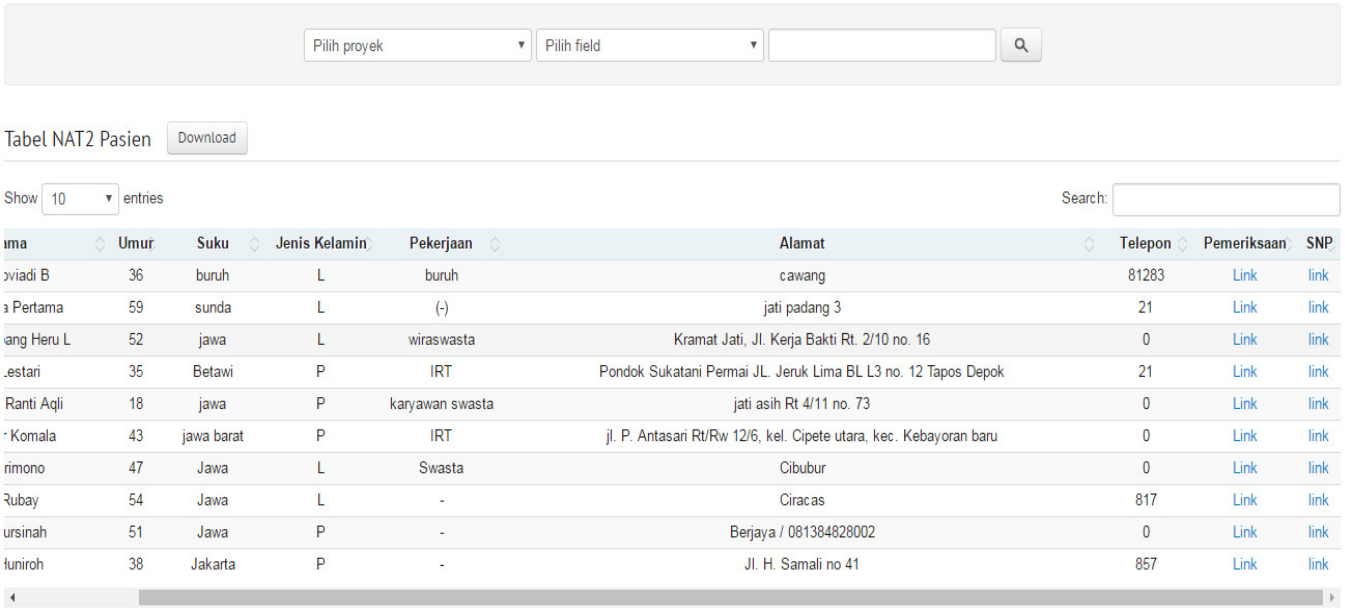


Figure 4: Genomic Medicine Research Group (Gmrg) Database Application Architecture

Search by a project name will display all patient information of each gene-associated disease belong to the project. Fig. 5 and 6 display all patient data in NAT2 and HLA database category, respectively. These figures display the detail information of patients' Electronic Medical Record (EMR) such as identity, jobs, ethnic, and their medical histories. The display of search result operation is structured into two levels hence it is easier for the user to track the more detail medication history of patients. GMRGDB browser is going to disclose those data when users choose the given link under SNP or Pemeriksaan link.



ıma	Umur	Suku	Jenis Kelamin	Pekerjaan	Alamat	Telepon	Pemeriksaan	SNP
ɔviadi B	36	buruh	L	buruh	cawang	81283	Link	link
ı Pertama	59	sunda	L	(-)	jati padang 3	21	Link	link
ıang Heru L	52	jawa	L	wiraswasta	Kramat Jati, Jl. Kerja Bakti Rt. 2/10 no. 16	0	Link	link
.estari	35	Betawi	P	IRT	Pondok Sukatani Permai JL. Jeruk Lima BL L3 no. 12 Tapos Depok	21	Link	link
Ranti Aqli	18	jawa	P	karyawan swasta	jati asih Rt 4/11 no. 73	0	Link	link
' Komala	43	jawa barat	P	IRT	jl. P. Antasari Rt/Rw 12/6, kel. Cipete utara, kec. Kebayoran baru	0	Link	link
ırimono	47	Jawa	L	Swasta	Cibubur	0	Link	link
Rubay	54	Jawa	L	-	Ciracas	817	Link	link
ursinah	51	Jawa	P	-	Berjaya / 081384828002	0	Link	link
ɬuniroh	38	Jakarta	P	-	Jl. H. Samali no 41	857	Link	link

Figure 5: Search Result Of NAT2 Data In GMRGDB

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Tabel hla_pasien Download

Show 10 entries Search:

ıur	Suku	Jenis Kelamin	Berat Badan	Tinggi Badan	Pekerjaan	Alamat	Merokok	No Rekam Medis	Diagnosis	Riwayat	Riwayat Keluarga	Terapi	Darah	Dna	Konsentrasi Dna	SNP
8	Betawi	P	0	0	-	Pekayon, Pasar Rebo	(-)	2013468442	-	-	-	-	?	?	0	link
5	-	L	0	0	-	-	(+) sudah 15 tahun berhenti	486582	Tb Paru dengan low intake	-	-	-	?	?	340	link
3	Betawi	L	0	0	-	-	(+) berhenti sudah 2 tahun	2013482842	-	-	-	-	?	?	402	link
0	bangka	P	0	0	IRT	Jl. Aup Barat no 2 rt/rw 2/6, kel. Jati padang, kec. Pasar minggu	(-)	394608	Tb paru	(-)	(-)	rifampisin 350,Isoniazid 300, pirazinamid 750, dan etambutol 500	?	?	269	link
6	Sunda	P	0	0	-	-	(-)	2013473521	Tb Paru dengan gangguan hati	-	-	-	?	?	164	link
ı	-	-	0	0	-	-	-	-	-	-	-	-	?	?	512	link

Figure 6: Search result of HLA data in GMRGDB

GMRGDB application enables users to filter the data based on the specific ethnic. Fig. 7 and 8 illustrate the results of searching for Betawi ethnic on NAT2 database and Javanese ethnic on HLA database, respectively. Through this feature, the researcher might swiftly have a broad illustration about which Indonesian ethnic that is sensitive in reacting to NAT2 enzyme after consuming Tuberculosis drugs, e.g. Isoniazid. The user may also fast to discover the ethnic inhabitants with HLA alleles that are diagnosed with more diseases. Besides search on ethnic, the user can find the number of patients that have been treated with such medicine, such as Rifampicin as illustrated in Fig. 9. In this figure, the researchers may draw valuable information by associating the medicines treatment to their diagnostic and family history. Fig. 10 presents the SNPs variation of each individual in N-acetyltransferase 2 gene that reacts to anti-TB drugs. This display exhibits alleles and genotyping information of NAT2 enzymes after reacting to the given medicines. The user can access this detail by clicking the given link under SNP column in the search result tables.

Pilih proyek Pilih field

Filter: Nat2 Pasien (suku) memiliki nilai 'Betawi'

Tabel NAT2 Pasien Download

Show 10 entries Search:

No	Nama	Umur	Suku	Jenis Kelamin	Pekerjaan	Alamat	Telepon	Pem
1	Siti Rohayati	38	Betawi	P	-	Pekayon, Pasar Rebo	0	
3	Tn. Ferry Kurniawan	33	Betawi	L	-	-	0	
7	Tn. Badaruddin	70	betawi	(-	SD	kuningan timur Rt.004/RW.003	0	
9	Tn. Rojali	63	betawi	L	pensiunan	kel. Rambutan Rt/Rw 13/2 no. 7	0	
10	Ny. Khairani	49	betawi	P	IRT	Rawa Bambu 1 Rt 6/6 no. 8 D jl. H	0	
14	Ny. Lestari	35	Betawi	P	IRT	Pondok Sukatani Permai JL. Jeruk Lima BL L3 no. 12 Tapos Depok	21	
22	Tn. Lugina	57	Betawi	L	-	Kampung Baru, Pnacoran Has RT 06 RW 06	0	
24	Tn. M. Daud	63	betawi	L	(-)	Jl. Menteng Atas rt/rw 9/5,menteng atas, setia budi	0	
28	Tn. Mahadi L	43	Betawi	L	(-)	jl. Merdeka, Ciracas	0	
33	Tn. M. Ali Asni	66	Betawi	L	supir	Jl. Kampung bali XXX/11 rt/rw 08/08 Kp. Bali, Tanah Abang	0	

Figure 7: Search result of Betawi ethnic on NAT2 database

Journal of Theoretical and Applied Information Technology
31st March 2017. Vol.95. No 6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

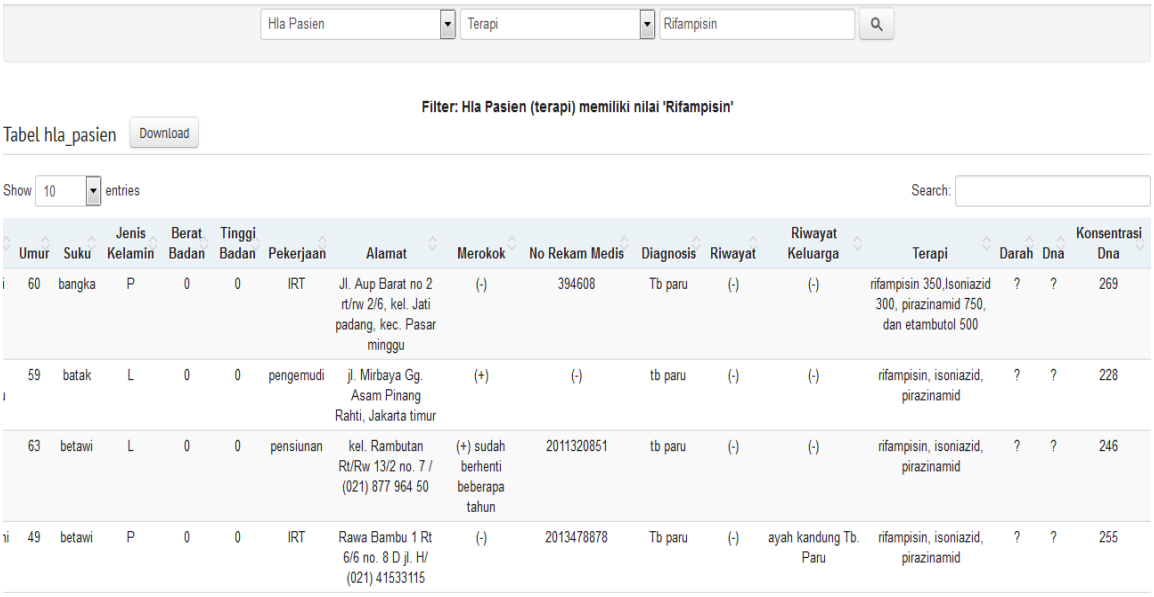
Filter: Hla Pasien (suku) memiliki nilai 'jawa'

Tabel hla_pasien Download

Show 10 entries Search:

nur	Suku	Jenis Kelamin	Berat Badan	Tinggi Badan	Pekerjaan	Alamat	Merokok	No Rekam Medis	Diagnosis	Riwayat	Riwayat Keluarga	Terapi	Darah	Dna	Konsentrasi Dna	SNP
52	jawa	L	0	0	wiraswasta	Kramat Jati, Jl. Kerja Bakti Rt. 2/10 no. 16	(+) berhenti merokok > 1 bln	2013483598	TB paru	(-)	(-)	rifampisin 450, isoniazid 300, pirazinamid 1000, dan etambutol 1000	?	?	346	link
18	jawa	P	0	0	karyawan swasta	jati asih Rt 4/11 no. 73	(-)	2013490032	TB paru dengan riwayat DIH	(-)	(-)	etambutol 750 dan streptomisin 750	?	?	169	link
47	Jawa	L	0	0	Swasta	Cibubur	(+) sudah 6 tahun	2013468711	TB Paru insuf Hepar	-	-	Rifampicin 450 INH 300 Ethambuthol 2 x 500 Streptomisin 1 x 1 gr	?	?	188	link
54	Jawa	L	0	0	-	Ciracas / 08177734395	(+)	2012392738	TB Paru Hepatitis OAT	-	-	-	?	?	4	link
51	Jawa	P	0	0	-	Berjaya / 081384828002	-	2013491849	DIH	-	-	-	?	?	29	link
52	jawa	L	0	0	(-)	Bambu Apus, Cipayung	(+) sudah berhenti 2bulan	2013501265	Tb paru	(-)	(-)	rifampisin 450, isoniazid 300, pirazinamid 1000, dan	?	?	263	link

Figure 8: Search result of Javanese ethnic on HLA database



Hla Pasien | Terapi | Rifampisin

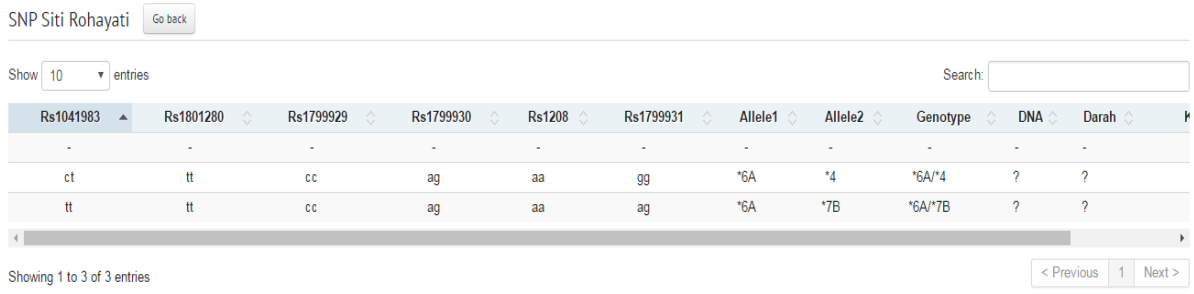
Filter: Hla Pasien (terapi) memiliki nilai 'Rifampisin'

Tabel hla_pasien Download

Show 10 entries Search:

Umur	Suku	Jenis Kelamin	Berat Badan	Tinggi Badan	Pekerjaan	Alamat	Merokok	No Rekam Medis	Diagnosis	Riwayat	Riwayat Keluarga	Terapi	Darah	Dna	Konsentrasi Dna
60	bangka	P	0	0	IRT	Jl. Aup Barat no 2 rt/rw 2/6, kel. Jati padang, kec. Pasar minggu	(-)	394608	Tb paru	(-)	(-)	rifampisin 350,Isoniazid 300, pirazinamid 750, dan etambutol 500	?	?	269
59	batak	L	0	0	pengemudi	jl. Mirbaya Gg. Asam Pinang Rahti, Jakarta timur	(+)	(-)	tb paru	(-)	(-)	rifampisin, isoniazid, pirazinamid	?	?	228
63	betawi	L	0	0	pensiunan	kel. Rambutan Rt/Rw 13/2 no. 7 / (021) 877 964 50	(+) sudah berhenti beberapa tahun	2011320851	tb paru	(-)	(-)	rifampisin, isoniazid, pirazinamid	?	?	246
49	betawi	P	0	0	IRT	Rawa Bambu 1 Rt 6/6 no. 8 D jl. H/ (021) 41533115	(-)	2013478878	Tb paru	(-)	ayah kandung Tb. Paru	rifampisin, isoniazid, pirazinamid	?	?	255

Figure 9: Genomic Medicine Research Group (Gmrg) Database Application Architecture



SNP Siti Rohayati Go back

Show 10 entries Search:

Rs1041983	Rs1801280	Rs1799929	Rs1799930	Rs1208	Rs1799931	Allele1	Allele2	Genotype	DNA	Darah
-	-	-	-	-	-	-	-	-	-	-
ct	tt	cc	ag	aa	gg	*6A	*4	*6A/*4	?	?
tt	tt	cc	ag	aa	ag	*6A	*7B	*6A/*7B	?	?

Showing 1 to 3 of 3 entries < Previous 1 Next >

Figure 10: Snps Information Of NAT2 Gene For One Individual

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5. CONCLUSIONS

We developed a modular architecture of Genomic Medicine Research Group Database (GMRGDB) to manage gene-associated diseases in Indonesia. This database covers most Indonesian ethnics, such as Javanese, Betawi, Sundanese, Melayu, and Batak. Currently, GMRGDB consists two database categories, N-acetyltransferase 2 (NAT2) and Human Leukocyte Antigen (HLA) group. These databases record the interconnection among diseases, genes, and reaction to drugs. GMRGDB browser provides the functions to view a number of samples for each database, find SNPs information, filter the data based on specified ethnic, and connect to other genomic browsers, e.g. NCBI website.

GMRGDB database will be continuously enriched with other gene-associated diseases. Leprae data for Indonesian population will be appended to the GMRGDB in the coming work. GMRGDB browser should provide the feature to automatically investigate the relationships between SNP and the given drugs. The associations may be visually presented by generating an interactive and informative dashboard with a befitting graphs or charts.

ACKNOWLEDGMENT

We are fully thankful to The Indonesian Directorate General of Higher Education (DIKTI) of the Ministry of Higher Education, Research and Technology, and also YARSI Foundation for providing grants to execute this research.

REFRENCES:

[1] Yuliwulandari R, Kashiwase K, Nakajima H, Uddin J, Susmiarsih TP, Sofro a. SM, et al. Polymorphisms of HLA genes in Western Javanese (Indonesia): Close affinities to Southeast Asian populations. Tissue Antigens. 2009;73(1):46–53.

[2] Yuliwulandari R, Sachrowardi Q, Nishida N, Takasu M, Batubara L, Susmiarsih TP, et al. Polymorphisms of promoter and coding regions of the arylamine N-acetyltransferase 2 (NAT2) gene in the Indonesian population: Proposal for a new nomenclature. J Hum Genet. 2008;53(3):201–9.

[3] De Santis D, Dinauer D, Duke J, Erlich H a, Holcomb CL, Lind C, et al. 16(th) IHIW : review of HLA typing by NGS. Int J Immunogenet [Internet]. 2013 Feb [cited 2015 Feb 17];40(1):72–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23302098

[4] Shiina T, Hosomichi K, Inoko H, Kulski JK. The HLA genomic loci map: expression, interaction, diversity and disease. J Hum Genet [Internet]. 2009 Jan [cited 2015 Jan 28];54(1):15–39. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19158813

[5] Sanchez-mazas A, Meyer D. The Relevance of HLA Sequencing in Population Genetics Studies. Journal of Immunology Research. 2014;12 pages. Available from: http://dx.doi.org/10.1155/2014/971818

[6] Butcher NJ, Sim E, Minchin RF. Pharmacogenetics of the arylamine N - acetyltransferases. Journal of pharmacogenomics. 2002;2:30–42.

[7] Huang YS. Recent progress in genetic variation and risk of antituberculosis drug-induced liver injury. J Chinese Med Assoc [Internet]. Elsevier Taiwan LLC and the Chinese Medical Association; 2014;77(4):169–73. Available from: http://dx.doi.org/10.1016/j.jcma.2014.01.010

[8] Yuliwulandari R, Susilowati RW, Wicaksono BD, Viyati K, Prayuni K, Razari I, et al. NAT2 variants are associated with drug-induced liver injury caused by anti-tuberculosis drugs in Indonesian patients with tuberculosis. J Hum Genet [Internet]. Nature Publishing Group; 2016;(January):1–5. Available from: http://www.nature.com/doifinder/10.1038/jhg.2016.10

[9] Sherry, S.T. et al. dbSNP: the NCBI database of genetic variation. Nucleic Acids Research, 2001, Vol. 29, No. 1, pp 308-311.

[10] Maglott, D. et al. Entrez Gene: gene-centered information at NCBI. Nucleic Acids Research, 2005, Vol. 33, Database issue, 54-58.

[11] NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. Nucleic Acids Research, 2013, Vol. 41, Database issue, 8-20.

[12] NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. Nucleic Acids Research, 2015, Vol. 43, Database issue, 6-17.

[13] Hirawa, M. et al. JSNP: a database of common gene variations in the Japanese population. Nucleic Acids Research, 2002, Vol 30, No.1, 158-162.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[14] Higasa, K. et al. Human genetic variation database, a reference database of genetic variations in the japanase population. Journal of Human Genetics, 2016, Vol. 61, 547-553.

[15] Halim-Fikri et al. The first Malay database toward the ethnic-spesific target molecular variation. BioMed Central Research Notes, 2015, Vol. 8:176.

[16] Irmert, F., Daum, M., and Meyer-Wegener, K. A new approach to modular database systems. Proceedings of the 2008 EDBT workshop on Software Engineering for tailor-made data management. France, 2008, 40-44.

[17] Kardiana, A. and Diana, N.E. Visualization of ontology-based data warehouse for malaria spread incidences using protégé. Jurnal Sistem Informasi, 2016, Vol. 12, No. 1, 10-16.

[18] Diana, N.E. and Kardiana, A. Comprehensive centralized-data warehouse for managing malaria cases. International Journal of Advanced Computer Scienc and Applications, 2015, Vol. 1, No. 6, 40-46.