

## **CHAPTER III**

### **RESEARCH METHOD**

#### **III. 1. Research aspects**

##### **III. 1. 1. Research field**

This research project is in the field of human genetics, mainly molecular genetics and ophthalmology.

##### **III. 1. 2. Research location**

Probands and their families were collected based on the medical record obtained from Dr. Kariadi Hospital, Central Java Eye Center, Semarang Eye Center and Low Vision Unit Yogyakarta. DNA isolation will be performed in the Center for Biomedical Research, Faculty of Medicine, Diponegoro University Semarang. Homozygosity mapping and sequencing analysis of candidate genes was performed in the Department of Human Genetics, Radboud University Nijmegen Medical Center (RUNMC), the Netherlands.

##### **III. 1. 3. Research period**

Sample collection and DNA isolation were performed within 6 months, followed by molecular genetic analysis in the next 12 months in 2010 - 2011.

### **III. 1. 4. Research design**

This study is a *descriptive, cross sectional* study

## **III. 2. Method**

### **III. 2. 1. Population**

Individuals with sporadic or familial RP were included in this study. The patients originated from several regions in Central Java with various ethnic backgrounds. Information about ethnical backgrounds of the patients were obtained by pedigree drawing of at least three family generations and information about the origin of the proband's ancestor.

### **III. 2. 2. Samples**

Patients who received a clinical diagnosis of non syndromic retinitis pigmentosa were included in this study. Parents and siblings were also included in this study in order to perform segregation analysis. From the previous study (13), 16 families consist of 36 affected and 29 unaffected members has been collected. Seven of these families had been solved by the previous study. In the present study, samples are obtained from 28 families with non syndromic RP, consist of 35 affected individuals and 40 unaffected family members. Blood samples of 180 random, unrelated, healthy individuals with no symptoms or family history of visual impairment were also collected as controls and used for

the determination of allele frequency in the general Indonesian population.

Informed consent obtained from each patient and control individual after explanation of the nature and possible consequences of the study in accordance with the tenets of the Declaration of Helsinki (Edinburgh, 2000; [www.wma.net](http://www.wma.net)).

### **III. 2. 2. 1. Inclusion criteria**

1. Probands who were clinically diagnosed with RP
2. Agree to join on this study and signing the consent form.
3. Patients with unilateral disease, nystagmus /wandering eye, exudative retinal detachment, retinal vasculitis
4. Patients who did not display any signs or symptoms of syndromic RP (retinitis pigmentosa with involvement of other organ, such as hearing problem and mental retardation) based on informations obtained from the history taking and clinical examination
5. Patients did not display any symptom or fundus appearance of retinitis pigmentosa sine pigmento

### III. 2. 2. 2. Operational Definition

No	Variable	Definition	Results / measurements	Scale
1.	Non Syndromic RP ( <b>dependent variable</b> )	is a group of inherited retinal degeneration disorder characterized by night blindness, progressive loss of peripheral vision and display a typical fundus appearance: bone spicule pigmentation, arterioles attenuation, waxy pallor of optic disc without any involvement of other organ.	Yes / No	Nominal
2.	Mutated gene ( <b>independent variables</b> )	Gene with a change in the of base pair in the DNA sequence	Yes / No	Nominal
3.	X-linked RP (xlRP)	RP that the defective genes segregate to the next generation on the X chromosome.	Yes / No	Nominal
4.	Autosomal dominant RP (adRP)	A condition when one mutated copy of the gene in each cell is sufficient for a person to be affected with RP	Yes/ No	Nominal
5.	Autosomal recessive RP (arRP)	A condition when two mutated copies of the gene are needed to cause RP phenotype in an affected person	Yes / No	Nominal
6.	Homozygosity mapping	A method for <u>mapping the human genome</u> , used to detect genes that cause disease only when both copies in an individual are mutated (i.e. the genes are homozygous, or the same in two allele). The regions determined from the SNP array data which has been analyzed using PLINK software.	Yes/No	Nominal

### III. 2. 2. 3. Subject selection

The probands were selected based on the medical records obtained from Dr. Kariadi Hospital, Central Java Eye Center, Semarang Eye Center and Low Vision Unit Yogyakarta. Subject were selected using *Non-probability, convenient sampling* method.

#### **III. 2. 2. 4. Clinical examination**

History taking and family pedigree were obtained from the participants and other family members. Basic ophthalmic examination, such as visual acuity by Snellen chart, color vision by Ishihara plate of the affected individuals performed by dr. Kentar Arimadyo, SpM followed by further detailed investigations, including direct ophthalmoscope funduscopy, Humphrey visual field analysis and fundus photography.

#### **III. 2. 2. 5. Sample collection**

After clinical diagnosis of RP established, 5 mL EDTA blood was drawn for DNA isolation from all affected and unaffected family members as well as control individuals.

#### **III. 2. 2. 6. Minimum sample required**

There are no minimal number of samples required, as this study performed in a is a descriptive study and not a population study.

### **III. 3. Laboratory procedures**

#### **III. 3. 1. DNA isolation**

Blood samples were collected from the affected and other available family members by vein puncture. Total genomic DNA were

extracted from peripheral leukocytes according to a standard salting-out protocol (Cebior manual book).

### **III. 3. 2. DNA concentration and Gel electrophoresis**

Before measuring the concentration make sure that the DNA quality is good by running the DNA by agarose gel electrophoresis using  $\lambda$  Hind III marker. The concentration of agarose 0.8 % and put 5  $\mu$ l loading buffer and 2  $\mu$ l DNA (70ng/ $\mu$ l) and set the electrophoresis on 50V for at least 2 hours. After DNA quality has been checked, DNA concentration was measured using the nanodrop and working solution with concentration of 70 ng/ $\mu$ l was made for sending SNP microarray samples. A DNA aliquot with concentration of 20 ng/ $\mu$ l was prepared for PCR working solution. Working solution was placed in eppendorf tube or sample plate, labelled for each sample and stored in the -20<sup>0</sup>C fridge.

### **III. 3. 3. SNP microarray and homozygosity mapping**

Genotyping was performed on Infinium Human Omni-Express 700K arrays (Illumina, San Diego, CA) containing approximately 700,000 single nucleotide polymorphisms. Array experiments are according to the protocol provided by manufacturer. Homozygous regions are determined by PLINK software (51), using a sliding window of 50 SNPs, 10 missing SNPs per window and allowing 2 heterozygous

SNPs. Homozygous regions larger than 3 Mb were included as homozygous regions.

Homozygous regions were ranked based on the size of the regions. SNP positions were derived from the UCSC Human Genome Browser build NCBI 36/ hg 18, March 2006. Gene candidates selection will ranked from the largest homozygous region and for the familial cases, in overlapping regions shared by siblings. Available published retinal dystrophy loci and genes data (RetNet-Retinal Information Network; <http://www.sph.uth.tmc.edu/retnet/>, provided in the public domain by the University of Texas Houston Health Science Center, Houston, TX) which provide data for the RP known genes was used to performed gene candidates selection within the homozygous regions.

Each homozygous regions may harbor many genes residing the homozygous regions. Prioritization of the most promising genes as the disease-causing genes is needed to determine which gene has to be tested first. Therefore, several available database were used for this prioritization which based on the known mutations, protein function, gene expression and the phenotype of knock out mice with defects in this gene (Table 2).

**Table 2. Overview of the databases that were used for candidate gene search**

PROVIDED DATA	DATABASE/ WEBSITE
<b>Known mutations</b>	<ul style="list-style-type: none"> <li>• <b><i>The Human Gene Mutation Database:</i></b> <a href="http://www.hgmd.cf.ac.uk/ac/index.php">http://www.hgmd.cf.ac.uk/ac/index.php</a></li> <li>• <b><i>Online Mendelian Inheritance in Man (OMIM):</i></b> <a href="http://www.ncbi.nlm.nih.gov/sites/entrez?db=omim">http://www.ncbi.nlm.nih.gov/sites/entrez?db=omim</a></li> <li>• <b><i>NextProt:</i></b> <a href="http://beta.nextprot.org/db/">http://beta.nextprot.org/db/</a></li> <li>• <b><i>RetNet (Retinal information Network):</i></b> <a href="http://www.sph.uth.tmc.edu/retnet/">http://www.sph.uth.tmc.edu/retnet/</a></li> </ul>
<b>Protein function</b>	<ul style="list-style-type: none"> <li>• <b><i>Uniprot:</i></b> <a href="http://www.uniprot.org/">http://www.uniprot.org/</a></li> <li>• <b><i>UCSC Genome Browser:</i></b> <a href="http://genome.ucsc.edu/">http://genome.ucsc.edu/</a></li> </ul>
<b>Gene expression</b>	<ul style="list-style-type: none"> <li>• <b><i>GeneCards:</i></b> <a href="http://www.genecards.org/">http://www.genecards.org/</a></li> <li>• <b><i>GeneHub-GEPIS From Gene Intergration to Expression Profiling:</i></b> <a href="http://www.cgl.ucsf.edu/Research/genentech/genehub-gepis/index.html">http://www.cgl.ucsf.edu/Research/genentech/genehub-gepis/index.html</a></li> <li>• <b><i>Tissue Distribution Databases:</i></b> <a href="http://genius.embnet.dkfz-heidelberg.de/menu/cgi-bin/srs/wgetz?-page+query5+-l+hs_tissue_distribution">http://genius.embnet.dkfz-heidelberg.de/menu/cgi-bin/srs/wgetz?-page+query5+-l+hs_tissue_distribution</a></li> <li>• <b><i>TiGER, Wilmer Eye Institute of Johns Hopkins University:</i></b> <a href="http://bioinfo.wilmer.jhu.edu/tiger/">http://bioinfo.wilmer.jhu.edu/tiger/</a></li> </ul>
<b>Knock-out mice phenotype</b>	<ul style="list-style-type: none"> <li>• <b><i>Mouse Genome Informatics Jackson Lab:</i></b> <a href="http://www.informatics.jax.org/">http://www.informatics.jax.org/</a></li> </ul>

### III. 3. 4. Primers and Conditions for PCR

In patients with known RP genes residing within the largest homozygous regions, amplification of all exons and intron–exon boundaries of these genes were performed. Primers for the candidate gene are designed using Primer3 software (52) and are listed together with the PCR conditions in appendix. Designing primers is essential for



PCR optimization and to get the correct product. Below are the several considerations to design a good PCR primers:

1. **Length:** The optimal length of PCR primers is 18-22 bp, enough for adequate specificity, and primers can bind easily to the template at the annealing period.
2. **Melting Temperature( $T_m$ ):** is the temperature at which half of the DNA duplex will dissociate becomes single stranded and indicates the duplex stability. Primers with melting temperatures in the range of 52-58 °C generally produce the best results.
3. **Annealing temperature:** The reaction temperature is lowered to 50–65 °C for 20–40 seconds allowing annealing of the primers to the single-stranded DNA template. Typically the annealing temperature is about 3-5 degrees Celsius below the  $T_m$  of the primers used.
4. **GC Content:** The GC content (the number of G's and C's in the primer as a percentage of the total bases) of primer should be 40-60%.
5. **GC Clamp:** The presence of G or C bases within the last five bases from the 3' end of primers (GC clamp) helps promote specific binding at the 3' end due to the stronger bonding of G and C bases. More than 3 G's or C's should be avoided in the last 5 bases at the 3' end of the primer.

There is no single set of PCR mix and conditions that can be applied to all PCR amplifications. Optimization has to be done for each exon of genes and the PCR components (concentration of each agent, enzyme, temperature, number of cycles and requirements of additional agent, such as DMSO, Q-Solution, different type of polymerase) need to be adjusted within proposed ranges for specific product. Control DNA was used to adjust the condition for PCR. If the reaction worked properly, patient's DNA can be amplified using this protocol with including the positive and negative control to know that the reaction is working properly and to know if there is any contamination in the PCR mix.

The following universal PCR mix and conditions can be used for the first set to start amplification. The protocol that could be used / set per reaction for PCR as follow: H<sub>2</sub>O 16,5 µl, PCR buffer 10x 2,5 µl, 50 mM MgCl<sub>2</sub> 1,5 µl, dNTPs 10 mM 0,5 µl, 10 pmol Primer forward 0,5 µl, 10 pmol Primer reverse 0,5 µl, polymerase 5U/ul 0,5 µl, DNA 2,5 µl, and the total amount of this mix is 25 µl.

The universal PCR condition is 94 °C for 3 minutes; 94 °C for 30 second , annealing temperature is 58 °C for 30 seconds and elongation is 72 °C for 30 seconds, repeat in 35 times of cycle and ended with 72 °C for 5 minutes.

### III. 3. 5. Sequencing analysis

Purification of PCR product using Nucleospin Plasmid Quick Pure columns (Macherey-Nagel) and analyzed in either sense or antisense direction using dye termination chemistry (Big Dye Terminator, ver.3, model 3730 DNA analyzer; Applied Biosystems, Inc.[ABI], Foster City, CA). Sequencing result were analyzed using contig-VNTI program which allign and compared the reference sequence with the sample sequence.

### III. 3. 6. Mutation analysis

In silico programs, including Polymorphism Phenotype (*PolyPhen*), Sorting Intolerant from Tolerant (*SIFT*) and *Align GVGD* were used to determine pathogenicity of missense mutation. These programs are able to calculate the potential functional and structural impacts based on Grantham scores (comparison of the differences in physical properties of the amino acids' side chains) and PhyloP scores (evolutionary conservation of the nucleotide). Project hope was also used to visualized the protein's crystal structure changes caused by mutation.

Pathogenicity of nonsense mutations can be analyzed by considering several aspects such as the degree of conservation of amino acid changes, is the mutation located in an important domain, in which exon this mutation is located, and how is the possibility of Nonsense mediated decay occurence. However, most of nonsense mutations are

considering to result in the production of truncated protein which lead to the disease phenotype. Therefore, if a nonsense mutations found, it is we can almost always sure that this is the disease-causing mutations. Yet, confirmation test by segregation analysis and frequency analysis in ethnically matched control are still needed to be performed to confirm this mutations.

#### **III.3.6.1 Confirmation of a novel mutation**

In a novel mutation, segregation analysis in the parents and siblings are needed to be performed by direct sequencing. Sequencing analysis result will display whether the disease segregate within the family. The mutation can be confirmed if it is found exclusively in affected individual but absence in the unaffected parents and siblings. By performing this method, the mode of inheritance can also be determined and risk of the next generation can be calculated.

To determine whether the variant is a common variant/polymorphism, minimum number of 300 allele from unaffected, ethnically matched controls has to be screened for the mutation using PCR-restriction fragment length polymorphism (PCR-RFLP) or ARMS PCR.

### **III. 4. Data Collection**

#### III. 4.1. Primary Data

RP patients including demographic data: date of birth, gender and three generations pedigree

#### III. 4.2. Secondary Data

Medical records from the reference hospital, including patients clinical data, fundus appearance, fundus photograph and visual field analysis result.

### **III. 5. Data analysis**

Data were analyzed with descriptive method and presented in tables and graphics. Family pedigree were presented as figures with specific information for each families including family's origin, degree of consanguinity, number of affected, is proband's family lived in an isolated region where the degree of marriages between relatives are high. RP symptoms, age of onset, ophthalmic examination and fundus appearance of all probands were summarized in a table. PCR optimization data are important for further studies, especially with the same genes candidate. Therefore, primer lists, primer sequences, PCR mix, PCR conditions and additional kit requirements are listed specifically for each exon of the genes which were presented in a different table for each gene tested. Mutation analysis result and the pathogenicity prediction were listed in a table and explained in details for each mutation. The

proportion of all identified mutations in Indonesian families was presented in a pie chart diagram.

### **III. 6. Research Ethics**

Informed consents were provided from the initial consent given by parent or caregiver at the time of diagnosis that allows the use of material left anonymously. Ethical clearance was provided by Health Research Ethical Committee of Faculty of Medicine, Diponegoro University and Dr. Kariadi General Hospital Semarang. Ethical clearance form is attached in the appendix.

### **III. 7. Research Limitations**

In the present study, segregation analysis is needed to confirm the mutation that found in the patients. Segregation analysis is a method to confirm a mutation by testing unaffected parents and siblings of affected individual to check whether the mutation is exclusively found only in the affected individual. Therefore, parents and siblings samples are important to obtained. Segregation analysis also important to determine the inheritance pattern of mutation. However, due to some reasons, some of parents and siblings does not agree to be involved in this study. In these type families, inheritance modes cannot be determined, and information about the risk factor of the next generation cannot be assessed.

### III. 8. Research Flow

