

**DETEKSI SINYAL SELEKSI ALAM PADA POPULASI PAPUA  
MENGUNAKAN DATA SNP *GENOME-WIDE***

**SKRIPSI**

Untuk memenuhi sebagian persyaratan  
guna memperoleh gelar Sarjana Sain



Disusun oleh:

Monika Meili Novita

M0416032

**PROGRAM STUDI BIOLOGI  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS SEBELAS MARET  
SURAKARTA  
2020**

**PENGESAHAN**

**SKRIPSI**

**DETEKSI SINYAL SELEKSI ALAM DI POPULASI PAPUA  
MENGUNAKAN DATA SNP *GENOME-WIDE***

Oleh:  
Monika Meili Novita  
M0416032

Telah dipertahankan di depan Tim Penguji  
pada tanggal 24 Juli 2020  
dan dinyatakan telah memenuhi syarat

Surakarta, 24 Agustus 2020

Menyetujui

Penguji I

Penguji II

Dr. Ari Susilowati, S.Si, M.Si  
NIP. 196904281997022006

Suratman, S.Si, M.Si  
NIP. 198007052002121002

Pembimbing I

Pembimbing II

Prof. Drs. Sutarno, M.Sc., Ph.D.  
NIP. 196008091986121001

Pradiptajati Kusuma, Ph.D.

Mengetahui  
Kepala Program Studi S1 Biologi



Dr. Ratna Setyaningsih, M.Si  
NIP. 19660714199903200

### PERNYATAAN

Dengan ini saya menyatakan bahwa skripsi ini adalah hasil penelitian saya sendiri dan tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu perguruan tinggi, serta tidak terdapat karya atau pendapat yang pernah ditulis, atau diterbitkan oleh orang lain, kecuali secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.  
Apabila di kemudian hari dapat ditemukan adanya unsur penjiplakan maka gelar kesarjanaan yang telah diperoleh dapat ditinjau dan atau dicabut.

Surakarta, 24 Juli 2020



Monika

(Monika Meili Novita)

(M0416032)

# DETEKSI SINYAL SELEKSI ALAM DI POPULASI PAPUA MENGUNAKAN DATA SNP *GENOME-WIDE*

MONIKA MEILI NOVITA

Program Studi Biologi Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Sebelas Maret Surakarta

## ABSTRAK

Sejak awal kedatangan manusia modern dari Afrika (*Out-of-Africa*) sekitar 70.000-50.000 tahun yang lalu, Indonesia telah menjadi daerah persimpangan untuk migrasi manusia. Gelombang awal migrasi *Out-of-Africa* menuju Indonesia merupakan leluhur langsung dari orang yang mendiami wilayah Papua saat ini. Pola hidup orang asli Papua sangat bergantung dari alam dengan aktivitas berburu-meramu dan hidup berpindah-pindah. Pada kondisi kesehatan dan penyakit, wilayah Papua merupakan wilayah endemis malaria, namun orang Papua dapat bertahan hidup di tengah tekanan parasit yang tinggi. Perbedaan kondisi geografis, leluhur, pola hidup dapat membentuk berbagai variasi genetik pada populasi Papua, salah satunya dapat dilihat dari data polimorfisme nukleotida tunggal (SNP) yang juga digunakan dalam penelitian ini. Studi ini dilakukan untuk mengetahui gen dan varian genetik terseleksi yang menjadi sinyal seleksi alam pada populasi Papua dengan menggunakan data SNP *genome-wide*. Penelitian ini menggunakan data genotipe *genome-wide* pada 445 individu dari 18 populasi di Asia Tenggara Dataran, Kepulauan Asia Tenggara, Melanesia dan Polinesia. Data varian genom dianalisis untuk mendapatkan profil demografi dari populasi dan divisualisasikan menggunakan *Principal Component Analysis* (PCA). Gen dan varian yang terseleksi pada populasi Papua dilihat melalui hasil analisis indeks fiksasi ( $F_{ST}$ ) dan *Population Branch Statistic* (PBS), kemudian *Gene Ontology* dianotasikan menggunakan DAVID. Hasil analisis  $F_{ST}$  dan PBS menunjukkan gen *ASAP1* sebagai sinyal seleksi terkuat yang berperan dalam adaptasi populasi Papua terkait sistem syaraf. Selain itu, hasil pengayaan *Gene Ontology* juga menunjukkan gen-gen yang berperan dalam sistem imun (*HLA-DQ1* dan *CD226*), dan juga terkait metabolisme alkohol (*EPHB1* dan *ADH4*). Gen-gen tersebut menjadi kandidat seleksi positif pada populasi Papua dan berkaitan dengan perbedaan pola hidup serta kondisi geografis/penyakit di wilayahnya.

Kata kunci : genetika populasi, indeks fiksasi ( $F_{ST}$ ), Papua, *Population Branch Statistic* (PBS), seleksi positif

# NATURAL SELECTION SIGNAL DETECTION IN PAPUAN POPULATION USING GENOME-WIDE SNP DATA

MONIKA MEILI NOVITA

Biology Study Program, Faculty of Mathematic and Natural Sciences,  
Sebelas Maret University, Surakarta

## ABSTRACT

Since the arrival of modern humans from Africa (Out-of-Africa) around 70,000-50,000 years ago, Indonesia has been an intersection for human migration. The initial wave of Out-of-Africa migration to Indonesia bore the ancestor of people who inhabit Papua today. The ways of life of indigenous Papuans are very dependent on nature since ancient times, by hunting and gathering. Moreover, the Papua region has been a malaria endemic area. Nevertheless, the Papuans have survived the unique lifestyle and the high parasitic pressure in the environment. The differences in geographical conditions, environmental pressures, and lifestyles form a genetic adaptation in the Papuan population, which can be seen from single nucleotide polymorphism (SNP) data and also used in this study. This study was conducted to find out the selected genes and genetic variants as natural selection signal in the Papuan population using Genome-wide SNP data. This study used genome-wide genotype data of 445 individuals from 18 populations in the Mainland Southeast Asia, the Island Southeast Asia, Melanesia and Polynesia. Genomic variant data were analyzed to obtain demographic profiles and population structure visualized using Principal Component Analysis. Selected genes and variants in the Papuan population were estimated by fixation index ( $F_{ST}$ ) and Population Branch Statistic (PBS) analyses. The Gene Ontology enrichment is annotated using DAVID. The results of  $F_{ST}$  and PBS analysis showed the *ASAP1* gene as the strongest selection signal in the Papuan population, which is related to the nervous system. Moreover, the results of Gene Ontology enrichment analysis showed terms related to immune system (*HLA-DQA1* and *CD226*), and alcohol metabolism (*EPHB1* and *ADH4*). These genes are the candidates of positive selection the Papuan population which suggest the correlation with their unique lifestyle and the geographical/disease conditions in the region.

Keywords: fixation index ( $F_{ST}$ ), Papuan, *Population Branch Statistic* (PBS), population genetics, positive selection

## MOTTO

*“Hidup adalah perjalanan, ketika kita berhenti berarti ada yang tidak beres”*

–Paus Fransiskus

*“Bersabarlah dengan segala hal, tapi terutama bersabarlah terhadap dirimu. Jangan hilangkan keberanian dalam mempertimbangkan ketidaksempurnaanmu, tapi mulailah untuk memperbaikinya, mulailah setiap hari dengan tugas yang baru.”*

–St. Fransiskus dari Sales

*commit to user*

## PERSEMBAHAN

*Karya ini spesial kupersembahkan untuk*



*Bapak dan Mama tercinta  
yang selalu menyanyangi dan mendukung dengan tulus dengan  
segala pengorbanan yang telah diberikan*

*Adik dan keluarga tersayang  
yang selalu menyertakan semangat dan dukungan penuh untuk diriku*

*Biologi 2016 "Biosolid"  
suka dan duka bersama kalian memberikan warnaselama masa perkuliahan*

*Seluruh mahasiswa biologi dimanapun berada,  
semoga penelitian ini dapat memberi manfaat untuk perkembangan  
ilmu pengetahuan di masa depan*

*commit to user*

## KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Kuasa atas rahmat dan karuniaNya sehingga penulis dapat menyelesaikan penelitian dan penyusunan skripsi yang berjudul: “Deteksi Sinyal Seleksi Alam pada Populasi Papua menggunakan Data SNP *Genome-wide*”. Penyusunan skripsi ini merupakan suatu syarat untuk memperoleh gelar kesarjanaan strata 1 (S1) pada Prodi Biologi, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Sebelas Maret Surakarta.

Selama proses penelitian maupun penyusunan skripsi ini, penulis telah mendapatkan banyak masukan, bantuan dan bimbingan dari berbagai pihak yang sangat berguna dan bermanfaat baik secara langsung maupun tidak langsung. Oleh karena itu pada kesempatan yang baik ini dengan berbesar hati penulis ingin mengucapkan terima kasih yang setulus dan sebesar-besarnya kepada :

1. Dr. Ratna Setyaningsih, M.Si. selaku Kepala Program Studi Biologi, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Sebelas Maret Surakarta yang telah memberikan izin dalam penelitian.
2. Prof. Amin Soebandrio, dr., Ph.D., Sp. MK(K), selaku Kepala Lembaga Biologi Molekuler Eijkman yang telah memberikan izin untuk bisa melakukan penelitian di Laboratorium Keagenaragaman Genom dan Penyakit.
3. Prof. Drs. Sutarno, M.Sc., Ph.D. selaku dosen pembimbing I yang telah memberikan bimbingan, saran dan motivasi selama penelitian sampai selesainya penyusunan skripsi.
4. Pradiptajati Kusuma, Ph.D. selaku dosen pembimbing II yang telah berkenan memberikan bimbingan, arahan, dukungan, saran dan motivasi selama penelitian sampai selesainya penyusunan skripsi ini. Penulis belajar banyak hal dan ilmu baru yang sangat berharga dari beliau.

*commit to user*



5. Dr. Ari Susilowati, M.Si. selaku dosen pembimbing akademik dan penguji I yang telah memberikan bimbingan, dukungan dan saran selama penyusunan skripsi.
6. Suratman, M.Si. selaku dosen penguji II yang telah memberikan bimbingan, dukungan dan saran selama penyusunan skripsi.
7. Kedua orang tua tercinta, Bapak dan Mama yang telah memberikan dukungan penuh dan doa untuk kelancaran penelitian hingga penyusunan skripsi.
8. Adik tersayang, Angela Dea Desideria, serta Dimitri Arioseno Ekaputra tersayang, yang selalu mendukung, memberikan keceriaan, semangat dan doa kepada penulis selama penelitian dan penyusunan skripsi.
9. Sahabat tersayang, Jedidi'ah Abel, Roro Long Puteri, Yessia Evita, Gia, Kashish Harjani, Elsabella Limawan, Lintang Respati, Mareta, Bagas, Natali yang telah memberikan semangat, motivasi, dukungan, saran selama penyusunan skripsi. Terima kasih juga telah menjadi teman diskusi dan cerita yang baik.
10. Teman-teman biologi angkatan 2016 "*Biosolid*" yang telah memberikan warna dan semangat selama perkuliahan.
11. Semua pihak yang tidak dapat disebutkan satu per satu yang telah memberikan bantuan dan semangatnya kepada penulis hingga akhir penyusunan skripsi ini.

Penulis juga menyampaikan permohonan maaf kepada semua pihak yang berkaitan dengan penelitian ini apabila terdapat hal-hal yang kurang berkenan di hati. Penulis juga menyadari bahwa penelitian ini masih memiliki kekurangan. Oleh karena itu, saran dan masukan sangat penulis harapkan demi perbaikan. Penulis berharap naskah skripsi ini dapat bermanfaat bagi pengembangan ilmu pengetahuan di masa depan.

Surakarta, 24 Juli 2020

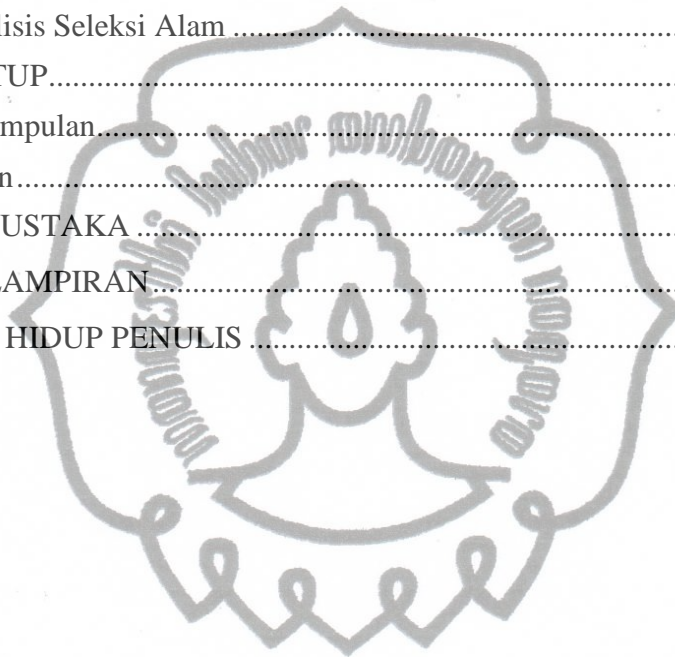
Penulis

*commit to user*

## DAFTAR ISI

HALAMAN JUDUL .....	i
PENGESAHAN .....	ii
PERNYATAAN.....	iii
ABSTRAK .....	iv
ABSTRACT.....	v
MOTTO .....	vi
PERSEMBAHAN.....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xii
DAFTAR GAMBAR .....	xiii
DAFTAR LAMPIRAN.....	xiv
DAFTAR SINGKATAN.....	xv
I. PENDAHULUAN.....	1
A. Latar Belakang .....	1
B. Rumusan Masalah.....	4
C. Tujuan Penelitian.....	4
D. Manfaat Penelitian.....	4
II. LANDASAN TEORI.....	5
A. Tinjauan Pustaka .....	5
1. Sejarah migrasi manusia menuju Indonesia .....	5
2. Seleksi Alam.....	8
3. Polimorfisme Nukleotida Tunggal .....	12
4. Teknik Indeks Fiksasi ( $F_{ST}$ ).....	14
5. Teknik <i>Population Branch Statistic</i> (PBS).....	15
B. Kerangka Berpikir .....	18
III. METODE PENELITIAN.....	20
A. Waktu dan Tempat Penelitian .....	20
B. Cara Kerja .....	20
1. Etik .....	20
2. Data kasar .....	20

3. <i>Quality Control</i> .....	22
4. Struktur Populasi .....	24
5. Analisis Seleksi positif .....	25
6. Analisis <i>Gene Ontology</i> .....	26
D. Analisis Data .....	26
IV. HASIL DAN PEMBAHASAN.....	27
A. Analisis Profil Demografi dan Struktur Populasi.....	27
B. Analisis Seleksi Alam .....	32
V. PENUTUP.....	44
A. Kesimpulan.....	44
B. Saran.....	44
DAFTAR PUSTAKA .....	45
DAFTAR LAMPIRAN.....	57
RIWAYAT HIDUP PENULIS .....	63



### DAFTAR TABEL

Tabel 1. Daftar Populasi dan Jumlah Individu data SNP Genome-Wide .....21

Tabel 2. Hasil pengayaan *Gene Ontology* 133 gen outlier hasil PBS menggunakan  
DAVID.....42



## DAFTAR GAMBAR

Gambar 1.a. Peta menunjukkan pembagian kepulauan Asia Tenggara.....	6
Gambar 1.b. Topografi ISEA saat ini setelah naiknya air laut.....	6
Gambar 2. Reskonstruksi genetik migrasi manusia di Kepulauan Asia Tenggara (ISEA).....	8
Gambar 3. Jenis seleksi alam dalam menentukan distribusi sifat .....	10
Gambar 4. Hasil pemindaian seluruh genom untuk seleksi positif.....	17
Gambar 5. Skema Kerangka Berpikir Penelitian .....	19
Gambar 6. Peta persebaran populasi pada data SNP Genome-Wide yang digunakan pada penelitian ini .....	21
Gambar 7. Plot PCA dari 445 individu di 18 populasi .....	29
Gambar 8. Visualisasi matriks dua dimensi (heatmap) dan dendrogram hubungan kekerabatan 18 populasi berdasarkan hasil perhitungan jarak genetik berpasangan ( $F_{ST}$ ) .....	31
Gambar 9.a. Plot manhattan berdasarkan hasil $F_{ST}$ antara populasi (a) Indonesia bagian timur dengan barat .....	34
Gambar 9.b. Papua dengan Indonesia bagian barat. Sumbu y menunjukkan nilai empiris dari $F_{ST}$ .....	34
Gambar 10. Plot manhattan berdasarkan hasil PBS antara tiga populasi. Sumbu y menunjukkan nilai empiris dari PBS, sedangkan sumbu x menunjukkan nomor kromosom autosom .....	36

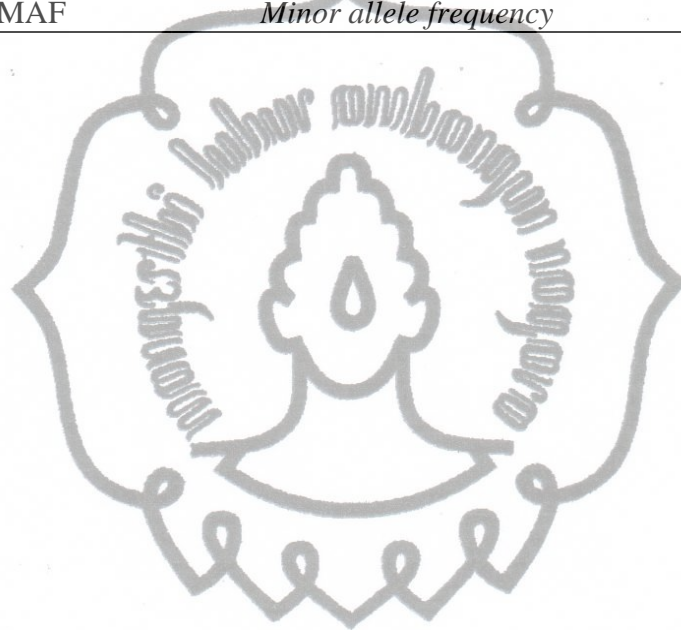
**DAFTAR LAMPIRAN**

Lampiran 1. Hasil pengayaan Gene Ontology 133 gen outlier hasil PBS  
menggunakan DAVID.....57  
Riwayat Hidup Penulis .....63



**DAFTAR SINGKATAN**

<b>Singkatan</b>	<b>Kepanjangan</b>
SNP	<i>Single nucleotide polymorphism</i>
PCA	<i>Principal component analysis</i>
$F_{ST}$	<i>Fixation Index</i>
PBS	<i>Population branch statistic</i>
LD	<i>Linkage disequilibrium</i>
MAF	<i>Minor allele frequency</i>



*commit to user*