

## **ANALISA BIOINFORMATIKA GEN E1 DAN E2 DARI VIRUS HEPATITIS C (HCV) GENOTIPE 1, 2, 3 DAN 6 SEBAGAI KANDIDAT VAKSIN VIRAL-LIKE PARTICLES (VLP)**

Henny Saraswati

Fakultas Ilmu-ilmu Kesehatan, Universitas Esa Unggul

Jl.Arjuna Utara No.9, Jakarta Barat

hennysaraswati@esaunggul.ac.id

### **Abstract**

*Hepatitis C virus (HCV) causes hepatitis disease. This disease is dangerous due to the large number of people with chronic hepatitis infection. Chronic hepatitis infection may progress to cirrhosis, liver cancer and death. The mortality rate due to liver cancer is quite large and occupies the top 10 causes of death in the world. HCV vaccine is necessary to prevent HCV transmission. The challenge of this vaccine development is the genotypes circulate in the world. In Indonesia, there are genotypes 1, 2, 3 and 6. HCV vaccine should be able to protect from most genotype. Therefore, the development of HCV vaccine takes a long time. One vaccine approach sounds promising is VLP (Viral-like Particles) vaccine. The E1 E2 protein on the surface of the virus is an excellent candidate for VLP vaccine material, because of its immunogenicity. In this study, we did bioinformatics study to analyze E1 and E2 genes from genotypes 1, 2, 3 and 6. These genes sequences are obtained from GenBank, and analyzed further using softwares such as BioEdit Sequence Alignment Editor, BLAST and BLAST Primer. From this study we can obtain E1-E2 gene consensus sequence and also conserved region of E1 and E2 gene sequence. The results of this study can be further used in the development of HCV VLP vaccine candidates.*

**Keywords :** VLP vaccine, E1 gene, E2 gene.

### **Abstrak**

Virus Hepatitis C (HCV) merupakan penyebab penyakit hepatitis. Penyakit ini berbahaya dikarenakan besarnya angka penderita infeksi hepatitis kronis. Infeksi hepatitis kronis dapat berlanjut menjadi sirosis dan kanker hati, hingga kematian. Angka kematian karena kanker hati cukup besar dan menempati 10 besar penyebab kematian tertinggi di dunia. Vaksin HCV sangat diperlukan untuk mencegah penularan infeksi. Kesulitan yang dihadapi adalah banyaknya genotipe HCV yang beredar. Di Indonesia sendiri terdapat genotipe 1, 2, 3 dan 6. Vaksin HCV harus dapat memberikan proteksi terhadap infeksi beberapa macam genotipe. Oleh karena itu, pengembangan vaksin HCV memerlukan waktu yang cukup lama. Salah satu pendekatan vaksin adalah dengan vaksin VLP (*Viral-like Particles*). Protein E1E2 yang terdapat pada permukaan virus merupakan kandidat yang sangat baik untuk dijadikan bahan vaksin VLP karena memiliki imunogenesitas yang tinggi. Dalam penelitian ini dilakukan analisa bio informatika terhadap gen E1 dan E2 dari genotipe 1, 2, 3 dan 6. Sekuen gen-gen ini diperoleh melalui Gen Bank dan dianalisa lebih lanjut menggunakan beberapa perangkat lunak seperti *Bio Edit Sequence Alignment Editor*, *BLAST* dan *Primer BLAST*. Berhasil didapatkan sekuen konsensus gen E1-E2 dan juga sekuen lestari dari gen E1 dan E2. Diharapkan hasil penelitian ini dapat digunakan lebih lanjut dalam pengembangan kandidat vaksin VLP HCV.

**Kata kunci :** Vaksin VLP, gen E1, gen E2.

### **Pendahuluan**

Virus Hepatitis C (HCV) merupakan penyebab penyakit sirosis dan kanker hati yang menjadi permasalahan kesehatan dunia. Secara global jumlah penderita hepatitis C kronis sekitar 130-150 juta orang. Sekitar 15-30% dari jumlah penderita kronis infeksi HCV ini akan mengalami sirosis (World Health Organization, 2016). Di Indonesia sendiri, sekitar 28 juta diketahui menderita infeksi hepatitis B dan C, 14 juta orang di antaranya berpotensi menjadi kronis. Sebanyak 1,4 juta orang

di antaranya berpotensi menderita sirosis (Pusdatin Kemenkes, 2014).

Pencegahan penularan virus HCV dapat dilakukan melalui vaksin. Namun, sayangnya vaksin untuk infeksi HCV saat ini belum tersedia. Penelitian pengembangan vaksin HCV telah dilakukan selama beberapa dekade. Berbagai pendekatan teknologi juga sudah dilakukan seperti pengembangan vaksin DNA, peptida sintetik, rekombinan protein, protein *core* dan protein *envelope*. Namun, beberapa tantangan yang ada

menghambat terbentuknya vaksin dari beberapa pendekatan ini. Salah satu tantangan yang dihadapi adalah variasi genetik HCV yang tinggi di seluruh dunia. Variasi genetik antara genotipe satu dengan yang lain sekitar 30%. Sehingga vaksin yang ada harus dapat menjangkau semua genotipe yang ada.

Salah satu pendekatan vaksin yang berpotensi untuk menjadi terapi profilaksis adalah vaksin VLP (*Virus-like Particles*). Vaksin VLP adalah vaksin yang berisi protein yang mirip dengan struktur virus asli serta memiliki imunogenesitas tinggi. Artinya, vaksin ini dapat menstimulasi respon imun yang mampu untuk mengeliminasi virus. Kelebihan dari VLP sebagai kandidat vaksin adalah lebih aman dan lebih murah dibandingkan dengan vaksin dengan pendekatan lain. Vaksin VLP terdiri dari protein struktural suatu pathogen. Pada virus Hepatitis C yang berpotensi untuk pembentukan vaksin VLP. Protein E1 dan E2 diketahui mampu menimbulkan respon imun humoral dan seluler yang adekuat pada beberapa hewan coba serta pada subyek penelitian orang sehat (tanpa infeksi HCV). Respon imun humoral (antibodi) yang ditimbulkan juga bersifat *cross-reactive*, artinya mampu memberikan proteksi terhadap beberapa infeksi HCV dari beberapa genotipe yang berbeda. Sehingga protein-protein ini sering digunakan sebagai kandidat vaksin profilaksis infeksi HCV.

Bio informatika menurut Kamus Oxford berarti adalah ilmu dalam mengumpulkan dan menganalisa data-data biologi seperti kode genetik. Terdapat juga arti lain, bahwa bio informatika adalah manajemen sistem informasi untuk biologi molekuler dan memiliki beberapa aplikasi fungsional (Luscombe, 2001). Bio informatika ini semakin berkembang dengan semakin banyaknya data-data molekuler organisme dari waktu ke waktu dan memerlukan suatu sistem untuk mengaturnya. Komputer menjadi jawaban dari permasalahan ini. Salah satu contohnya adalah proyek genom manusia yang selesai pada tahun 2003, menghasilkan data-data genom manusia, dan sekarang sering digunakan dalam bidang kesehatan seperti melihat keterkaitan gen terhadap penyakit. Aplikasi lain dari bio informatika ini cukup banyak, seperti dalam memprediksi fungsi gen dan protein, prediksi struktur protein, terapi gen, hubungan obat dengan gen, pembuatan obat baru, pengembangan vaksin, dll. Dalam pengembangan vaksin salah satu contohnya adalah pengembangan vaksin DNA untuk malaria. Dimana dilakukan pencarian data base gen-gen yang berpotensi untuk dikembangkan menjadi vaksin (Shuaibu, 2010).

Penelitian ini bertujuan untuk menganalisa gen E1 dan E2 dari virus HCV yang berasal dari genotipe 1, 2, 3 dan 6 dengan pendekatan

bioinformatika,. Diharapkan dari hasil penelitian ini dapat diketahui sekuen-sekuen lestari (*conserved*) pada gen-gen E1 dan E2 HCV dari genotipe 1, 2, 3 dan 6. Hasil penelitian ini juga dapat digunakan dalam pengembangan vaksin VLP HCV yang sangat dibutuhkan di Indonesia dan dunia pada umumnya.

## **Metode Penelitian**

### **Mendapatkan Database Gen e1 dan e2 HCV dari Genotipe 1, 2, 3 dan 6**

Database gen E1 dan E2 dari genotipe 1, 2, 3 dan 6 virus Hepatitis C diakses dari pusat data informasi genomik seperti pada *National Center for Biotechnology Information* (NCBI) yang berpusat di Amerika Serikat dan melalui *Virus Pathogen Resources* (ViPR) yang mengintegrasikan data-data virus patogen dari beberapa sumber, seperti GenBank, UniProt, Immune Epitop Database, Protein Data Bank. Data DNA HCV yang dipilah dari virus-virus HCV dengan genotipe 1, 2, 3 dan 6, yang berasal dari daerah Asia, virus tersebut menginfeksi manusia dan berasal dari hasil isolasi < 10 tahun. Data-data genome ini kemudian diunduh dalam bentuk FASTA.

### **Proses Pensejajaran (Alignment) Gen E1 dan E2 Hcv dari Genotipe 1, 2, 3 dan 6**

Proses ini dilakukan dengan bantuan perangkat lunak *Bioedit Sequence Alignment Editor ver.7.0.5.3* (North Carolina State University, AS). Untuk setiap genotipe, dilakukan pensejajaran dari 5 (lima) sekuen-sekuen gen E1 dan E2, sehingga didapatkan sekuen konsensus untuk setiap genotipe. Pada proses pensejajaran ini akan terlihat pada daerah mana suatu sekuen terdapat lestari pada semua genotipe dan pada daerah mana yang tidak.

### **Pensejajaran Sekuen Konsensus Gen E1 dan E2 dengan Seluruh Data Sekuen Hcv Genotipe 1, 2, 3 dan 6**

Proses pensejajaran kali ini dilakukan pada sekuen konsensus gen E1 dan E2 yang berhasil didapatkan dengan seluruh data genom HCV genotipe 1, 2, 3 dan 6 yang ada di GenBank. Proses ini dilakukan dengan perangkat lunak BLASTn (*Basic Local Alignment Search Tools-Nucleotide*) yang terdapat pada situs NCBI pada alamat [www.blast.ncbi.nlm.nih.gov](http://www.blast.ncbi.nlm.nih.gov). dari HCV. Sekuen konsensus dibandingkan untuk setiap genotipe, baik genotipe 1, 2, 3 dan 6. Hasilnya adalah tingkat kemiripan atau homologi dari sekuen konsensus dengan data genom HCV untuk genotipe 1, 2, 3 dan 6 yang ada. Tingkat kemiripan yang akan digunakan adalah 90%.

## Penentuan Daerah Lestari dari Sekuen Konsensus Gen E1 dan E2

Penentuan daerah lestari pada sekuen gen E1 dan E2 merupakan kelanjutan dari hasil pensejajaran sekuen dengan BLAST. Hasil pensejajaran sekuen diunduh dalam bentuk file FASTA kemudian dilakukan proses pencarian daerah lestari dengan bantuan perangkat lunak *BioEdit Sequence Alignment Editor*. Panjang basa daerah lestari adalah > 10 basa dengan tingkat homologi 100%.

## Desain Primer Untuk Konstruksi Plasmid Rekombinan Kandidat Vaksin VLP HCV

Hasil akhir dari pensejajaran sebelumnya berupa sekuen konsensus digunakan sebagai dasar pembuatan primer. Primer yang didesain harus dapat digunakan dalam perbanyakan gen. Syarat primer yang akan didesain memiliki panjang 18-24 pb baik forward dan reverse, dengan suhu  $T_m$  sekitar 55-65°C dan akan menghasilkan produk PCR sepanjang 576 untuk gen E1 dan 1089 pb untuk gen E2. Beberapa kriteria lain seperti primer harus tidak berikatan dengan pasangan primer lain pada sisi 3' dan tidak membentuk struktur *hairpin*. Untuk melakukan desain primer maka digunakan perangkat lunak Primer BLAST (*Basic Local Alignment Search Tool*) dari NCBI melalui <https://www.ncbi.nlm.nih.gov/tools/primer-blast/>.

## Pembahasan

### Pencarian Database Gen E1 dan E2 HCV dari Genotipe 1, 2, 3 dan 6

Setelah dilakukan pencarian pada situs ViPR didapatkan 21 data genom HCV sesuai dengan kriteria yang diinginkan. Data-data genom HCV ini berasal dari genotipe 1, 2, 3 dan 6. Sebagian besar virus ini diisolasi di negara-negara Asia, seperti Cina, Jepang, Pakistan, Nepal, Malaysia, Vietnam, Thailand dan Australia. Data-data ini dipilih karena sekuen genom yang beredar di Asia memiliki kemiripan yang tinggi dengan yang beredar di Indonesia. Dalam penelitian ini hanya 3 sekuen yang berasal dari Amerika Serikat, Kanada dan Swiss seperti data sekuen gen E1 dan E2 dari genotipe 1 HCV, dikarenakan belum banyak data dari genotipe ini yang berasal dari Asia. Meskipun demikian, genotipe 1 adalah genotipe yang paling banyak terdapat di seluruh dunia. Kemungkinan besar genom dari genotipe ini tidak memiliki variasi yang terlalu besar di belahan dunia. Virus HCV yang datanya digunakan dalam penelitian ini menginfeksi manusia (memiliki inang manusia). Data-data lengkap dari genom yang didapatkan dapat dilihat pada Tabel 1.

Tabel 1

Data-data genom HCV yang digunakan dalam penentuan sekuen E1 dan E2 dari genotipe 1, 2, 3 dan 6

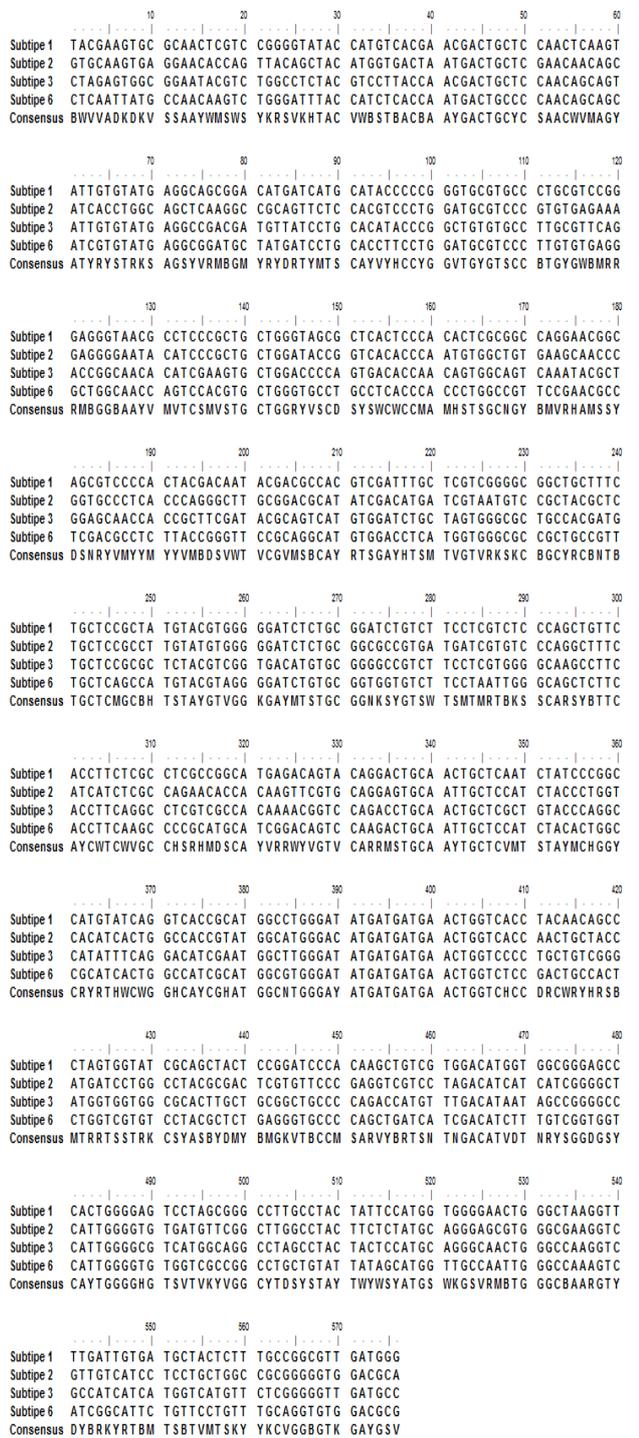
No	Strain	Genotipe	Nomor Akses GenBank	Tempat Isolasi
1	QC165	1c	KJ439767	Kanada
2	H4 HCV-	1b	GU451220	Cina
3	1a/CH/BID-V271/2006	1c	EU482858	Swiss
4	P1A01	1b	KX767012	Amerika Serikat
5	ZS542	2f	KC844042	Cina
6	ZS260	2b	KC844048	Cina
7	ZS623	2a	KC844043	Cina
8	J6CF/JFH2.1	2a	AB690461	Jepang
9	gz52540	1b	KC844051	Cina
10	SH37	3b	JQ065709	Cina
11	I	3a	KM587622	Australia
12	PK-1	3	GU294484	Pakistan
13	NE274	3d	KJ470619	Nepal
14	NE145	3e	KJ470618	Nepal
15	D	2b	KM587617	Australia
16	10MYKJ032	6n	KC191671	Malaysia
17	6_TV453	6	KJ567648	Vietnam
18	6-TV469	6	KJ567646	Vietnam
19	C-0782	6f	KM504110	Thailand
20	BL120	6a	KJ678744	Cina
21	6-TV443	6	KJ567649	Vietnam

Untuk setiap genotipe, dipilih 5-6 data sekuen genom sebagai rujukan. Hal ini dilakukan untuk menghindari kesulitan pembuatan sekuen konsensus setelah proses pensejajaran sekuen gen E1 dan E2. Sekuen pengkode gen E1 dan E2 didapatkan dari genom yang sama untuk menghindari variabilitas kedua gen ini. Tahun isolasi genom HCV yang digunakan dipilih dari tahun 2006-2012. Terdapat beberapa data genom yang tidak diketahui tahun isolasinya dari virus. Hal ini disebabkan karena peneliti yang melakukan proses isolasi DNA dan mengunggahnya di GenBank tidak mencantumkan kapan mereka melakukannya. Akan tetapi, data-data genom yang tidak diketahui tahun dilakukannya isolasi ini telah dipublikasikan di jurnal-jurnal ilmiah, pada tahun 2007-2016 (Lu, *et al*, 2007; Lu, *et al*, 2014; Li, *et al*, 2014; Lu, *et al*, 2016).

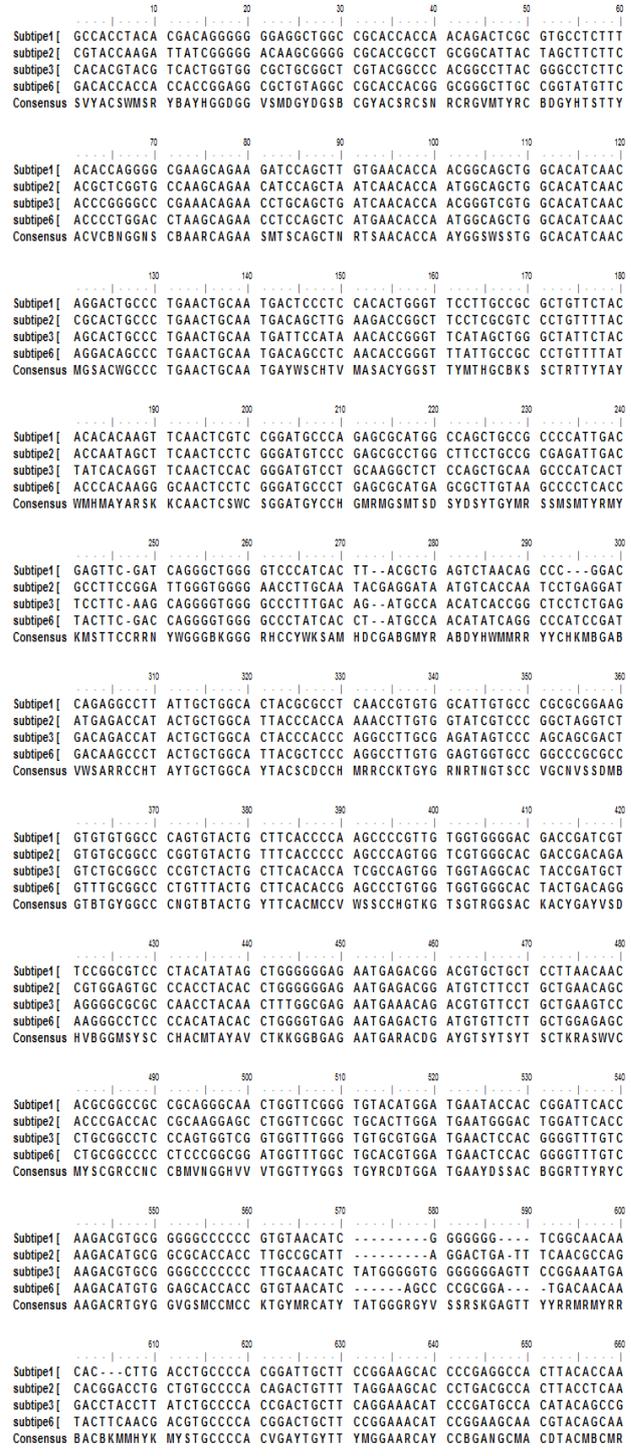
### Proses Pensejajaran (*Alignment*) Database Gen E1 Dan E2 Hcv dari Genotipe 1, 2, 3 dan 6

Penjajaran sekuen E1 dan E2 untuk genotipe 1 dilakukan terhadap 5 data gen E1, yaitu dari data strain QC165, H4, HCV-1a/CH/BID-V271/2006, P1A01 dan gz52540. Untuk genotipe 2, dilakukan pensejajaran dari strain ZS542, ZS260, ZS623, J6CF/JFH2.1 dan D. Cara yang sama juga dilakukan untuk genotipe 3, yaitu terhadap strain SH37, I, PK-1, NE274 dan NE145. Yang terakhir adalah pensejajaran untuk genotipe 6 yang dilakukan pada strain-strain berikut, yaitu 10MYKJ032, 6\_TV453,

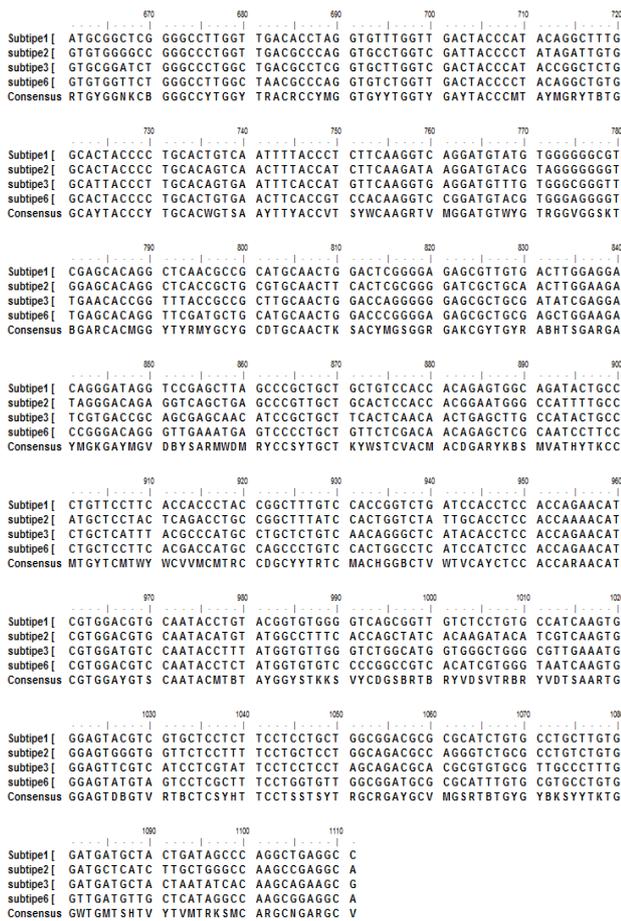
6-TV469, C-0782, BL120 dan 6-TV443. Dari pensejajaran untuk setiap genotipe ini di dapatkan sekuen konsensus. Kemudian sekuen konsensus untuk setiap genotipe ini disejajarkan kembali sehingga didapatkan sekuen konsensus untuk gen E1 dari genotipe 1, 2, 3 dan 6. Hasil dari pensejajaran ini dapat dilihat pada Gambar 1.



Hal yang sama juga dilakukan pada gen E2. Hasil pensejajaran dan sekuen konsensusnya dapat dilihat pada gambar 2.



Gambar 1.  
Hasilpensejajaransekuen-sekuen gen E1 darigenotipe 1, 2, 3 dan 6



Gambar 2

Hasil pensejajaran sekuen-sekuen gen E2 dari genotipe 1, 2, 3 dan 6

Dari hasil pensejajaran gen E1 dan E2 setiap genotipe ini terlihat adanya variasi di beberapa titik atau posisi. Variasi yang didapatkan untuk gen E1 ini berkisar 60% dan gen E2 sebesar 50%. Variasi yang cukup besar ini dikarenakan perbandingan dilakukan pada 4 genotipe yang berbeda. Tetapi terdapat juga daerah-daerah yang lestari. Variasi sekuen genetik pada HCV dengan genotipe yang sama adalah sekitar 15%, sedangkan variasi antar genotipe berkisar antara 30-35% (Messina et al., 2015; Smith et al., 2014). Variasi sekuen gen E1 dan E2 kemungkinan disebabkan karena proses adaptasi virus terhadap respon imunitas tubuh organisme inang, dalam hal ini adalah manusia. Protein E1 dan E2 merupakan glikoprotein yang terdapat pada bagian terluar dari virus dan diketahui berperan penting dalam proses masuknya (fusi) virus ke dalam sel inang. Dikarenakan posisi protein E1 dan E2 yang terletak di luar sel, maka akan dikenali oleh antibodi, baik antibodi monoklonal maupun antibodi netralisasi. Beberapa penelitian menemukan bahwa pada beberapa penderita hepatitis kronik memiliki titer antibodi netralisasi yang tinggi terhadap protein E1 dan E2. Oleh karena itu, virus HCV berupaya

untuk melakukan upaya penghindaran dari respon imun antibodi ini. Oleh karena itu, virus melakukan beberapa variasi gen E1 dan E2 agar tidak dikenali oleh respon imun tubuh manusia. Variasi sekuen gen ini banyak terdapat pada gen E2 yang juga terlihat pada proses pensejajaran kali ini. Protein E2 merupakan protein yang memiliki situs HVR1 (Hypervariable Region 1) dan HVR2 yang sangat bervariasi. Hal ini dibuktikan dengan beberapa penemuan bahwa HVR1 banyak terdapat pada strain-strain virus yang mampu bertahan dari serangan respon imun manusia.

### Pensejajaran Sekuen Konsensus Gen E1 dan E2 dengan Seluruh Data Sekuen HCV Genotipe 1, 2, 3 dan 6

Perbandingan dilakukan antara sekuen konsensus gen E1 dengan keseluruhan database genotipe 1, sebanyak 100 database genotipe 1 HCV (Gambar 3, tidak semua data dapat ditampilkan).

Sequences producing significant alignments:

Select: All None Selected (0)

Description	Max score	Total score	Query cover	E value	Ident	Accession
Hepatitis C virus subtype 10 isolate H260_10 E1 protein gene partial cds	493	493	90%	5e-139	70%	KU074491.1
Hepatitis C virus subtype 10 clone C-1 polyprotein gene partial cds	484	484	90%	3e-136	70%	EU057403.1
Hepatitis C virus subtype 10 gene for polyprotein partial cds strain: K1aE2L-1	473	473	90%	5e-133	70%	AB482228.1
Hepatitis C virus subtype 10 isolate V4540_10 E1 protein gene partial cds	471	471	90%	2e-132	70%	KU074491.1
Hepatitis C virus subtype 10 isolate V10_10 polyprotein gene partial cds	470	470	90%	6e-132	70%	KU073981.1
Hepatitis C virus subtype 10 gene for polyprotein partial cds strain: K1aE2L-9	470	470	90%	6e-132	70%	AB482225.1
Hepatitis C virus subtype 10 gene for polyprotein partial cds strain: K1aE2L-7	470	470	90%	6e-132	70%	AB482224.1
Hepatitis C virus subtype 10 gene for polyprotein partial cds strain: K1aE2L-5	470	470	90%	6e-132	70%	AB482223.1
Hepatitis C virus subtype 10 gene for polyprotein partial cds strain: K1aE2L-5	470	470	90%	6e-132	70%	AB482222.1
Hepatitis C virus subtype 10 gene for polyprotein partial cds strain: K1aE2L-4	470	470	90%	6e-132	70%	AB482221.1
Hepatitis C virus subtype 10 gene for polyprotein partial cds strain: K1aE2L-3	470	470	90%	6e-132	70%	AB482220.1
Hepatitis C virus isolate HCV-H2C-H2C-V092002 complete genome	470	470	90%	6e-132	70%	EU155307.2
Hepatitis C virus isolate HCV-H2C-H2C-V042002 complete genome	466	466	90%	7e-131	70%	EU155309.2
Hepatitis C virus isolate BHA_91_1577_3 polyprotein gene partial cds	464	464	90%	2e-130	70%	HU113698.1
Hepatitis C virus subtype 10 gene for polyprotein partial cds strain: K1aE2L-3	464	464	90%	2e-130	70%	AB482220.1
Hepatitis C virus isolate HCV-H2C-H2C-V142003 complete genome	464	464	90%	2e-130	70%	EU155308.2
Hepatitis C virus isolate HCV-H2C-H2C-V142004 complete genome	464	464	90%	2e-130	70%	EU155310.2
Hepatitis C virus subtype 10 isolate E11140_10 E1 protein gene partial cds	462	462	90%	6e-130	70%	KU073927.1
Hepatitis C virus subtype 10 isolate PK1140_10 E1 protein gene partial cds	461	461	90%	3e-129	70%	KU073928.1
Hepatitis C virus subtype 10 isolate PK1140_10 E1 protein gene partial cds	461	461	90%	3e-129	70%	KU073941.1

Gambar 3

Proses perbandingan sekuen gen E1 denganseluruh database HCV genotipe 1 (100 genom HCV genotipe 1, tidaksemuadapatditampilkan)

Perlu dicermati dari hasil ini adalah nilai *E-value* (Expect value) yang didapatkan. Nilai *E-value* adalah jumlah berapa kali suatu sekuen spesifik didapatkan sama dengan sekuen database apabila seseorang melakukan pencarian dalam suatu database. Dalam hal ini, semakin kecil nilai *E-value*

yang didapatkan dalam perbandingan, maka kemiripan antara sekuen-sekuen yang dibandingkan semakin besar. Selain itu didapatkan nilai % *identity*, dimana *identity* adalah kemiripan antara sekuen-sekuen yang diperbandingkan, memiliki data sekuen yang sama pada daerah yang sama. Nilai E-value untuk genotipe 1 yang tertinggi adalah  $5 \times 10^{-139}$  dengan % *identity* adalah 79%, artinya bahwa sekuen konsensus yang dibuat untuk gen E1 memiliki kemiripan yang tinggi dengan gen E1 dari virus-virus HCV genotipe 1 yang ada. Sehingga berpotensi digunakan sebagai *template* dalam pengembangan pembuatan vaksin VLP.

Untuk genotipe 2, 3 dan 6 tingkat kemiripan yang didapatkan berkisar antara 65-77% (Gambar 4a-c, tidak semua data dapat ditampilkan)

a)

Description	Max score	Total score	Query cover	E value	Ident	Accession
Hepatitis C virus genotype 2 gene for polyprotein, partial cds	410	410	98%	2e-115	76%	D13406.1
Hepatitis C virus subtype 2a isolate G2aK1, complete genome	316	316	93%	4e-87	73%	AF189003.1
Hepatitis C virus subtype 2a strain NC2a-1, complete genome	316	316	93%	4e-87	73%	AF230401.1
Hepatitis C virus subtype 2a isolate Z5523 polyprotein gene, complete cds	313	313	85%	5e-86	74%	KC344943.1

b)

Description	Max score	Total score	Query cover	E value	Ident	Accession
Hepatitis C virus subtype 3 isolate patient 289 polyprotein gene, partial cds	432	460	98%	1e-121	77%	GJ265200.1
Hepatitis C virus genotype 3 isolate P11-Sect1-3-S24 polyprotein gene, partial cds	420	420	90%	2e-120	77%	K0004552.1
Hepatitis C virus genotype 3 isolate P11-Picorna-3-S24 polyprotein gene, partial cds	420	420	90%	2e-120	77%	K0004540.1
Hepatitis C virus genotype 3 isolate P11-Picorna-7-S24 polyprotein gene, partial cds	420	420	90%	2e-120	77%	K0004547.1

c)

Description	Max score	Total score	Query cover	E value	Ident	Accession
Hepatitis C virus subtype 6a isolate CC216, complete genome	300	300	98%	2e-92	72%	EF424628.1
Hepatitis C virus subtype 6a isolate CC39, complete genome	277	305	97%	2e-75	70%	EF424625.1
Hepatitis C virus subtype 6a isolate CC227, complete genome	271	271	98%	6e-74	70%	EF424627.1
Hepatitis C virus subtype 6f isolate C-0044, complete genome	268	268	97%	1e-72	71%	DQ357301.1

Gambar 4.

Proses perbandingan sekuen gen E1 dengan seluruh database HCV (a) genotipe 2, (b) genotipe 3 dan (c) genotipe 6. Terdapat 100 genom yang diperbandingkan (tidak semua dapat ditampilkan)

Hasil yang didapatkan memperlihatkan adanya nilai *E-value* untuk gen E2 genotipe 1 sebesar 0 dan % *identity* sebesar 81% (Gambar 5.). Ini artinya sekuen konsensus yang dibuat memiliki kemiripan sebesar 81% dengan gen E2 pada virus-virus HCV genotipe 1 yang beredar di seluruh dunia (data tidak dapat ditampilkan semua).

Description	Max score	Total score	Query cover	E value	Ident	Accession
Hepatitis C virus subtype 1a isolate H7c polyprotein gene, partial cds	1007	1007	99%	0.0	81%	KJ327892.1
Hepatitis C virus subtype 1a isolate HCV-1a/USBC-1440208, complete genome	902	902	99%	0.0	80%	EU403065.1
Hepatitis C virus subtype 1a isolate G5254 polyprotein gene, complete cds	900	900	97%	0.0	81%	KC344951.1
Hepatitis C virus subtype 1a isolate HCV-1a/USBC-1721199, complete genome	900	900	99%	0.0	80%	EU155204.2
Hepatitis C virus subtype 1a isolate A208 core C208 polyprotein gene, partial cds	976	976	98%	0.0	80%	JN849751.1

Gambar 5

Proses perbandingan sekuen gen E2 dengan seluruh database HCV genotipe 1 (100 genom HCV genotipe 1, tidak semua dapat ditampilkan)

Hasil yang didapat juga memperlihatkan adanya kemiripan yang tinggi (77-80%) antara sekuen konsensus gen E2 dengan sekuen gen inidari virus HCV genotipe 2, 3 dan 6 yang ada dunia (Gambar 6a-c).

a)

Description	Max score	Total score	Query cover	E value	Ident	Accession
Hepatitis C virus genotype 2 gene for polyprotein, partial cds	951	951	99%	0.0	80%	D13406.1
Hepatitis C virus subtype 2a genome RNA, complete genome, strain 1607-17-1	823	823	99%	0.0	77%	AF890451.1
Hepatitis C virus subtype 2a isolate Z5523 polyprotein gene, complete cds	803	803	96%	0.0	77%	KC344943.1
Hepatitis C virus subtype 2a isolate NC2aK9, complete genome	792	792	96%	0.0	77%	AF189005.1
Hepatitis C virus subtype 2a isolate G2aK3, complete genome	705	705	96%	0.0	76%	AF189004.1

b)

Description	Max score	Total score	Query cover	E value	Ident	Accession
Hepatitis C virus S23 polyprotein gene, complete cds	770	770	94%	0.0	77%	GJ304261.1
Hepatitis C virus subtype 3a clone H16-E13 envelope protein 2 (E2) gene, partial cds	770	770	94%	0.0	77%	GQ266761.1
Hepatitis C virus subtype 3a clone H16-E28 envelope protein 2 (E2) gene, partial cds	770	770	94%	0.0	77%	GQ266661.1
Hepatitis C virus subtype 3a clone H16-E14 envelope protein 2 (E2) gene, partial cds	774	774	94%	0.0	77%	GQ266771.1
Hepatitis C virus subtype 3a clone H16-E12 envelope protein 2 (E2) gene, partial cds	774	774	94%	0.0	77%	GQ266751.1

c)

Sequences producing significant alignments:

Select: All None Selected 0/7

Description	Max score	Total score	Query cover	E value	Ident	Accession
Hepatitis C virus subtype 6 isolate C-1044, complete genome	749	749	97%	0.0	76%	DQ835780.1
Hepatitis C virus subtype 6 isolate C-1046, complete genome	733	733	92%	0.0	76%	DQ835784.1
Hepatitis C virus subtype 6 isolate QZ210, complete genome	729	729	96%	0.0	75%	EF424620.1
Hepatitis C virus subtype 6 isolate KM42, complete genome	729	729	94%	0.0	76%	AF107952.1
Hepatitis C virus subtype 6 isolate KM42, complete genome	729	729	94%	0.0	76%	DQ278894.1

Gambar 6

Proses perbandingan sekuen gen E2 dengan seluruh database HCV (a) genotipe 2, (b) genotipe 3 dan (c) genotipe 6. Terdapat 100 genom yang diperbandingkan (tidak semua dapat ditampilkan)

### Penentuan Daerah Lestari dari Sekuen Konsensus Gen E1 dan E2

Penentuan daerah lestari untuk sekuen konsensus gen E1 dan E2 diperlukan untuk mengetahui daerah-daerah mana saja yang tidak banyak memiliki variasi sekuen gen. Daerah lestari adalah daerah-daerah yang memiliki kemiripan sekuen yang tinggi. Sekuen-sekuen lestari ini dapat bermanfaat dalam pengembangan vaksin HCV dengan pendekatan yang berbeda-beda, pendekatan VLP, DNA sintetik maupun pendekatan lain. Proses pencarian dilakukan antara sekuen konsensus dengan seluruh sekuen gen E1 HCV yang ada pada database GenBank. Hal ini dibantu dengan perangkat lunak *BioEdit Sequence Alignment Editor*.

Dihasilkan beberapa daerah lestari baik pada gen E1 dan E2. Pada gen E1 virus HCV genotipe 1 terdapat 2 daerah lestari, yaitu pada posisi basa ke-33 hingga 53 dan basa ke-385-401 (Gambar 7). Untuk genotipe 2 terdapat 4 daerah lestari, yaitu posisi basa ke-426-442, 538-545, 562-572 dan 574-581 (Gambar 8). Untuk genotipe 3 terdapat 1 daerah lestari, yaitu pada posisi basa ke-451-464 (Gambar 9). Yang terakhir adalah pada genotipe 6 pada posisi basa ke-211-220, 223-230 dan 389-396 (Gambar 10).

Region 1: Position 33 to 53	Region 2: Position 385 to 401
Consensus:	401
33	Consensus:
GTCACGAACGACTGCTCCAAC	385
53	TGGGATATGATGATGAA
	401

Gambar 7

Hasil analisa daerah lestari gen E1 pada genotipe 1

Region 1: Position 426 to 442	Region 2: Position 538 to 545
Consensus:	Consensus:
426 ATGATGATGAACTGGTC 442	538 CACTGGGG 545
Region 3: Position 562 to 572	Region 4: Position 574 to 581
Consensus:	Consensus:
562 GCCTACTTCTC 572	574 ATGCAGGG 581

Gambar 8

Hasil analisa daerah lestari gen E1 pada genotipe 2

Region 1: Position 451 to 464
Consensus:
451 CCCAGACCTTGTT 464

Gambar 9

Hasil analisa daerah lestari gen E1 pada genotipe 3

Region 1: Position 211 to 220	Region 2: Position 223 to 230
Consensus:	Consensus:
211 GGGTCCGCA 220	223 CATGTGGA 230

Region 3: Position 389 to 396
Consensus:
389 TCCATCTA 396

Gambar 10

Hasil analisa daerah lestari gen E1 pada genotipe 6

Gen E2 juga memiliki daerah lestari meskipun terdapat sekuen HVR1 dan HVR2. Pada genotipe 1, daerah lestari yang didapatkan cukup banyak, yaitu terdapat pada 11 posisi, antara lain posisi (1) basa ke-140-160, (2) basa ke-260-280, (3) basa ke-424-443, (4) basa ke-480-494, (5) basa ke-520-536, (6) basa ke-556-572, (7) basa ke-638-657, (8) basa ke-731-747, (9) basa ke-890-905, (10) basa ke-974-991 dan (11) basa ke-1083-1105 (Gambar 11). Untuk genotipe 2, terdapat 2 daerah lestari yaitu pada posisi basa ke-104-122 (Gambar 12), genotipe 3 pada posisi basa ke-228-244 (Gambar 13) dan genotipe 6 pada posisi basa ke-341-348, 406-413, 558-566, 949-956 dan 1191-1198 (Gambar 14).

Region 1: Position 140 to 160	Region 2: Position 260 to 280
Consensus:	Consensus:
140	260
CTGAACTGCAATGACTCCCTC	CAGGGGTGGGGTCCCATCACT
160	280
Region 3: Position 424 to 443	Region 4: Position 480 to 494
Consensus:	Consensus:
424	480 GACGGACGTGCTGCT 494
GTGGTGGGGACGACCGATCG	
443	
Region 5: Position 520 to 536	Region 6: Position 556 to 572
Consensus:	Consensus:
520 GGCAACTGGTTCGGCTG 536	556 GGGTTCACCAAGACGTG
	572
Region 7: Position 638 to 657	Region 8: Position 731 to 747
Consensus:	Consensus:
638 TGCTTCGGGAAGCACCCGGA	731 TGGCACTACCCTGCAC
657	747
Region 9: Position 890 to 905	Region 10: Position 974 to 991
Consensus:	Consensus:
890 ACAACAGAG-TGGCAG 905	974 CGTGGACGTGCAATACCT
	991
Region 11: Position 1083 to 1105	
Consensus:	
1083	
GCCTGCTGTGGATGATGCTGCT	
1105	

Gambar 11

Hasil analisa daerah lestari gen E2 pada genotipe 1

Region 1: Position 104 to 122	Region 2: Position 474 to 492
Consensus: 104 CCAATGGCAGCTGGCACAT 122	Consensus: 474 CCTGGGGGAGAATGAGAC 492

Gambar 12.

Hasil analisa daerah lestari gen E2 pada genotipe 2

Region 1: Position 228 to 244	Region 2: Position 1166 to 1185
Consensus: 228 AGCTGCAAGCCCATCAC 244	Consensus: 1166 ATCGAGGATCGTGACCCGAG 1185

Gambar 13.

Hasil analisa daerah lestari gen E2 pada genotipe 3

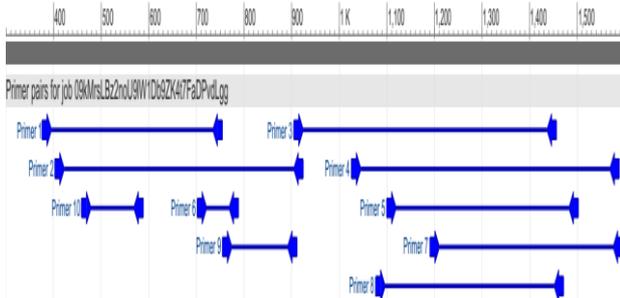
Region 1: Position 341 to 348	Region 2: Position 406 to 413
Consensus: 341 TGCTGGCA 348	Consensus: 406 TACTGCTT 413
Region 3: Position 558 to 566	Region 4: Position 949 to 956
Consensus: 558 TGGATGAAC 566	Consensus: 949 GAGCACAG 956
Region 5: Position 1191 to 1198	
Consensus: 1191 AAGTGGGA 1198	

Gambar 14.

Hasil analisa daerah lestari gen E2 pada genotipe 6

### Desain Primer untuk Konstruksi Plasmid Rekombinan Kandidat Vaksin VLP HCV

Penelitian ini dilakukan untuk mendapatkan pasangan primer untuk dapat menggandakan gen target E1-E2 secara utuh. Hasil yang didapatkan adalah sebagai berikut, sekuen konsensus gen E1-E2 memiliki panjang 1672 pb. Terdapat 10 pasang primer yang didapatkan dengan primer BLAST, memiliki daerah target masing-masing. Posisi pasangan primer E1-E2 dapat dilihat pada Gambar 15.



Gambar 15

Posisi relatif pasangan-pasangan primer terhadap gen E1-E2

Kesepuluh pasang primer ini rata-rata memiliki suhu Tm antara 55-60°C. Detil sekuen, posisi relatif primer, suhu Tm dan karakteristik lain dapat dilihat pada Tabel 2.

Tabel 2  
Detil karakteristik pasangan-pasangan primer terhadap gen E1 dan E2

Primer	Posisi	Sekuen (5'-3')	Panjang basa (basake)	Tm (°C)	GC%	3'-comple ntary
1	Forward	CGCATGGCCTGGGATATG	20 bp (376)	60,03	55	3
	Reverse	AAAACAGGGCGGCAATG	20 bp (754)	59,97	50	0
2	Forward	TGGTCATTTACTGCCGCC	20 bp (403)	60,03	50	2
	Reverse	CGGGACAATACCACAAG	20 bp (924)	60,04	55	3
3	Forward	AGCCTTGTGGTATTGTCC	20 bp (905)	60,04	55	2
	Reverse	TATGGCCAGCTCTGTTGT	20 bp (1456)	60,04	55	1
4	Forward	TGAGACGGATGTGTCTCT	20 bp (1026)	60,04	55	3
	Reverse	GTACTCCCACCTTGATGGC	20 bp (1588)	60,11	60	4
5	Forward	CCACGGGATTCGTC AAGA	20 bp (1100)	60,04	55	2
	Reverse	TGAGACCAAGTGGACAAA	20 bp (1502)	59,89	55	1
6	Forward	TGCCCTGAAC TCGAATGA	20 bp (702)	59,89	50	1
	Reverse	TCAGGACATCCCGAGGAG	20 bp (788)	59,96	55	2
7	Forward	TCAGGAAACATCCCGATG	20 bp (1191)	60,11	55	2
	Reverse	GGACGACG TACTCCC ACT	20 bp (1595)	60,11	50	2
8	Forward	GTTCCGGCTGTACGTGGAT	20 bp (1077)	60,11	55	1
	Reverse	GAAGGAGCAGGGCAGTA	20 bp (1471)	60,18	60	0
9	Forward	ACACCCACAGGTTCAACT	20 bp (755)	59,82	55	1
	Reverse	CAAGGCTTGGGTGCGTAA	20 bp (911)	59,83	55	1
10	Forward	CTTAGACATCATGGCCGG	20 bp (459)	60,25	60	2
	Reverse	GATGTAGGTGTC CGCGTC	20 bp (588)	60,11	55	0

Pasangan-pasangan primer ini kemudian mengalami perbandingan kembali (Re-BLAST) untuk mengetahui kecocokan primer dengan yang dimaksud serta untuk melihat apakah primer-primer ini dapat mengenali sekuen dari gen-gen lain (tidak spesifik). Apabila primer-primer ini tidak dapat secara spesifik mengenali gen E1 dan E2, maka sebaiknya tidak digunakan. Hal ini dikarenakan primer tadi akan mengenali gen lain dan menghasilkan perbanyakan gen yang tidak diinginkan. Hasil Re-BLAST dapat dilihat pada Tabel 3.

Tabel 3  
Hasil Re-BLAST Pasangan-pasangan Primer yang Berhasil Didapatkan

Primer	Nomor aksesi	Sekuen	Tingkat kesamaan (%)	E-value
Primer 1	KJ470506	CGCATGGCCTGGGATATGAT	100	0,28
	JQ470506	AAAACAGGGCGGCAATGAAC	100	0,011
	KY620832	TGGTCATTTACTGCCGCCAT	100	39
Primer 2	AB622121	CGGGACAATACCACAAGGCT	100	8e-04
	AB622121	AGCCTTGTGGTATTGTCCCG	100	8e-04
Primer 3	KM587623	TATGGCCAGCTCTGTTGTGG	95	0,18
	KY620850	TGAGACGGATGTGTCTCTGC	100	7e-04
Primer 4	KM007429	GTACTCCCACCTTGATGGCCC	100	7e-04
	KY620527	CCACGGGATTCGTC AAGACA	100	0,057
Primer 5	KY930698	TGAGACCAAGTGGACAAAGCC	100	0,012
	AY545952	TGCCCTGAAC TCGAATGACT	100	7e-04
Primer 6	KJ470622	TCAGGACATCCCGAGGAGTT	100	0,045
	KY620800	TCAGGAAACATCCCGATGCC	100	7e-04
Primer 7	KX621451	GGACGACG TACTCCC ACTTG	100	7e-04
	KY930697	GTTCCGGCTGTACGTGGATGA	100	6e-04
Primer 8	KY620797	GAAGGAGCAGGGCAGTATGG	100	6e-04
	KJ678753	ACACCCACAGGTTCAACTCC	100	7e-04
Primer 9	KT735649	CAAGGCTTGGGTGCGTAATG	100	0,042
	KY735649	CTTAGACATCATGGCCGGGG	100	0,19
Primer 10	EU925552	GATGTAGGTGTC CGCGTCAA	100	0,010

Semua pasangan primer yang didapatkan bisa dipilih untuk perbanyak gen E1-E2, namun apabila dilihat kembali pada karakteristik yang dimiliki, maka pasangan primer 1, 3, 7 dan 9 akan dipilih untuk proses perbanyak gen. Pasangan primer ini dipilih dikarenakan posisi relatifnya yang saling berdekatan dan tumpang tindih untuk dapat digunakan dalam perbanyak gen E1-E2 secara utuh. Selain itu primer-primer ini juga sedikit sekali memiliki sekuen yang komplemen terhadap dirinya sendiri pada daerah 3'. Hal ini penting karena primer tidak boleh terlalu banyak memiliki situs komplemen dengan dirinya sendiri yang dapat mengganggu proses penempelan primer ke gen target. Hasil akhirnya adalah tidak banyaknya gen target yang terduplikasi.

### **Kesimpulan**

Sesuai dengan tujuan penelitian kali ini, maka dari penelitian ini berhasil didapatkan sekuen konsensus gen E1-E2 yang berasal dari genotipe 1, 2, 3 dan 6 yang memiliki panjang 1672 pb. Sekuen ini menjadi dasar untuk desain pasangan primer dan 10 primer diperoleh. X sekuen primer telah dikonfirmasi menggunakan BLAST dan X sekuen primer memiliki tingkat homologi tinggi (> 90%). Hasil penelitian ini dapat dijadikan dasar dalam penelitian lebih lanjut, baik jangka pendek maupun panjang. Untuk jangka pendek akan digunakan dalam proses pengujian sekuen pasangan primer yang berhasil didapatkan.

### **Daftar Pustaka**

- Ashfaq,U.A, T.Javed, S.Rehman, Z.Nawaz dan S.Riazuddin. (2011). *An overview of HCV molecular biology, replication and immune responses*. Virol J. 8:161.
- Asselah, T, N. Boyer, D. Saadoun, M. Martinot-Peignoux dan P. Marcellin. (2016). *Direct-Acting Antivirals for The Treatment of Hepatitis C Virus Infection: Optimizing Current IFN-free Treatment and Future Perspective*. Liver Int. 36 (Suppl. S1): 47-57.
- Bartosch, B, J. Dubuisson dan F.C Cosset. (2003). *Infectious Hepatitis C Virus Pseudoparticles Containing Functional E1-E2 Envelope Protein Complexes*. J. Exp. Med. 197(5): 633-642.
- Bassett, S.E, D.L Thomas, K.M. Brasky, dan R.E Lanford. (1999). *Viral Persistence, Antibody to E1 and E2, and Hypervariable Region 1 Sequence stability in Hepatitis C Virus-inoculated Chimpanzees*. J. Virol. 73(2): 1118-1126.
- Beaumont, E dan P. Roingard. (2013). *Prospect for Prophylactic Hepatitis C Vaccines Based on Virus Like Particles*. Hum. Vaccin. Immunother. 9(5): 1112-1118.
- Beaumont, E, E. Roch, L. Chopin, dan P. Roingard. (2016). *Hepatitis C Virus E1 and E2 Proteins Used as Separate Immunogens Induce Neutralizing Antibodies with Additive Properties*. PLoS ONE. 1-15.
- Deng, K, R.Liu, H.Rao, D.Jiang, J.Wang, X.Xie, L.Wei. (2015). *Antibodies Targetting Novel Neutralizing Epitopes of Hepatitis C Virus Glycoprotein Preclude Genotype 2 Virus Infection*. PloS ONE 10(9): 1-17.
- Dubuisson, J. (2007). *Hepatitis C virus proteins*. World J Gastroenterol 2007 May 7; 13(17): 2406-2415.
- Frey, S.E, M.Houghton, S.Coates, S.Abrignani, D.Chien, D.Rosa, P.Pileri, R.Ray, A.M.Di Bisceglie, P.Rinella, H.Hill, M.C. Wolff, V.Schultze, J.H. Han, B.Scharschmidt, R.B. Belshe. (2010). *Safety and immunogenicity ofHCV E1E2vaccineadjuvantedwithMF59ad ministered to healthy adults*. Vaccine 28 (2010) 6367–6373.
- Garrone, P, A.C.Fluckiger, P.E. Mangeot, E. Gauthier, P.Dupeyrot-Lacas, J.Mancip, A. Cangialosi, I.Du Chéné, R. LeGrand, I.Mangeot, D.Lavillette, B.Bellier, F.L Cosset, F.Tangy, D.Klatzmann, C.Dalba. (2011). *A Prime-Boost Strategy Using Virus-Like Particles Pseudotyped for HCV Proteins Triggers Broadly Neutralizing Antibodies in Macaques*. Sci. Transl. Med., 3(94):94ra71.
- Halliday, J., P.Klenerman dan E. Barnes. (2011). *Vaccination for Hepatitis C Virus: Closing an Evasive Target*. Expert Rev. Vaccine. 10(5): 659-672.
- InfoDATIN. (2014). *Situasi dan Analisis Hepatitis*. Kementerian Kesehatan Republik Indonesia.
- Kushnir, N, S.Streatfield, V.Yusibov. (2012). *Virus-like particles as a highly efficient vaccine platform: Diversity of targets and*

*production systems and advances in clinical development.* Vaccine. 31(1):58-83.

- Li, C, V.H. Pham, K. Abe, L. Lu. (2014). *Nine Additional Complete Genome Sequences of HCV Genotype 6 from Vietnam Including New Subtypes 6xb and 6xc.* Virology. 468-470: 172-177.
- Li, L, C. Li, Y. Fu, F. Gao, O.G. Pybus, K. Abe, H. Okamoto, C.H. Hagedorn, D. Murphy. (2007). *Complete Genomes of Hepatitis C Virus (HCV) Subtypes 6c, 6l, 6o, 6p and 6q: Completion of a Full Panel of Genomes for HCV Genotype 6.* J. Gen. Virol. 88(5):1519-1525.
- Lu, L, T. Wu, L. Xiong, C. Li, M.H. Nguyen, D.G. Murphy. (2016). *Analysis of HCV-6 Isolates Among Asian-born Immigrants in North America Reveals Their High Genetic Diversity and a New Subtype.* Virology. 492: 25-31.
- Lu, L, Y. Yu, D.G. Murphy. (2014). *Full-length genomes of 16 hepatitis C virus genotype 1 isolates representing subtypes 1c, 1d, 1e, 1g, 1h, 1i, 1j and 1k, and two new subtypes 1m and 1n, and four unclassified variants reveal ancestral relationships among subtypes.* J. Gen. Virol. 95 : 1479-1487.
- Martínez-Donato, G, B. Piniella, D. Aguilar, S. Olivera, A. Pérez, Y. Castañedo, L. Alvarez-Lajonchere, S. Dueñas-Carrera, J.W. Lee, N. Burr, M. Gonzalez-Miro, B.H.A. Rehm. (2016). *Protective T-Cell and Antibody Immune Responses against Hepatitis C Virus Achieved Using a Biopolyester-Bead-Based Vaccine Delivery System.* Clin Vaccine Immunol 23:370–378.
- Messina, J.P, I. Humphreys, A. Flaxman, A. Brown, G.S. Cooke, O.G. Pybus dan E. Barnes. (2015). *Global Distribution and Prevalence of Hepatitis C Virus.* Hepatology. 61: 77-87.