

Reka Bentuk DNA Mengguna Sistem Imun Buatan Objektif Tunggal

M.Daman Huri
Mohd Zakree Ahmad Nazri
Azuraliza Abu Bakar
Tribasuki Kurniawan

ABSTRAK

Jujukan struktur asid deoksiribonukleik (DNA) yang tepat penting bagi komputeran DNA agar proses pengiraan berasaskan DNA menghasil keputusan yang berkesan. Untuk menghasil jujukan DNA yang baik, ramai penyelidik berusaha mengurangi kesilapan pengiraan dalam komputeran DNA dengan memberi tumpuan kepada pembangunan kaedah mereka bentuk jujukan DNA yang baik. Kajian ini membangun kaedah mereka bentuk jujukan DNA berasaskan sistem imun buatan (SIB) bagi mengoptimimum reka bentuk jujukan DNA. Teknik Algoritma Pilihan Negatif (APN) dan Algoritma Pilihan Klonal dihibrid dan satu algoritma baharu dihasil bagi mereka bentuk jujukan DNA. Algoritma ini dipanggil Algoritma berasaskan Pemilihan Negatif dan Klonal bagi mereka bentuk DNA atau PNK-D. Secara asasnya, terdapat empat kekangan yang berkesan terhadap gerak balas jujukan DNA iaitu; $H_{sukatan}$, keserupaan, kesinambungan, selekoh-tajam dan dua kekangan lain untuk mengekal keseragaman dalam turutan iaitu kandungan-GC dan suhu leleh (T_m). Nilai keempat-empat kekangan ini diagregasi sebagai satu objektif tunggal. Hasil keputusan kajian ini dibanding dengan kajian lepas yang mengguna kaedah pengoptimimuman koloni semut (ACO) bagi menyelesaikan permasalahan reka bentuk jujukan DNA. Hasil perbandingan menunjukkan jujukan DNA dihasil AIS lebih baik daripada jujukan yang dihasil mengguna kaedah ACO.

Kata kunci: Sistem imun buatan, reka bentuk jujukan DNA, komputeran DNA, komputeran lembut, pemilihan negatif

ABSTRACT

The right sequence structure of a deoxyribonucleic acid (DNA) is very important for DNA computing to obtain effective results. In order to generate good DNA sequences, many researchers try to reduce the calculation error in DNA computing by focusing on developing new methods for design in good DNA sequence. This study has developed a method based on artificial immune systems (AIS) to optimize the design of DNA sequence. Negative Selection Algorithm (APN) and Clonal Selection algorithm were hybridized. This algorithm is called Negative and Clonal Selection for the design of DNA or PNK-D. Basically there are four constraints that affects the response of the DNA sequence; $H_{measure}$, similarity, continuity, hairpin and two other constraints to maintain uniformity in the sequence, which is Gc content and melting temperature (T_m). In this study, values of the four constraints are aggregated as a single objective. The results of this study was compared with previous studies that uses antcolonyoptimization (ACO) to solve the design problems of the DNA sequence. The results showed that the DNA sequence produced AIS is better than the sequences produced by ACO rules.

Keywords: Artificial immune system, DNA sequence design, DNA computing, soft computing, negative selection

PENGENALAN

Komputeran DNA atau komputeran molekular adalah sekumpulan bebenang atau untaian DNA buatan yang dipilih khas. Komputeran DNA adalah kaedah komputeran mengguna DNA buatan seperti rangkaian neural buatan yang berasaskan metafora rangkaian neural otak manusia. Hasil kombinasi bebenang DNA ini menghasil penyelesaian kepada suatu masalah. Misalnya, menyelesaikan masalah pengoptimimuman dengan pantas berbanding komputeran konvensional. Kepantasan komputeran DNA berpunca daripada penggunaan konsep DNA yang membolehkan pengkomputeran selari secara besar-besaran. Sekiranya komputeran konvensional hendak melakukan komputeran selari besar-besaran, maka perkakasan yang

banyak diperlu sekali gus dapat meningkat kos operasi. Bagi komputeran DNA, komputeran selari besar-besaran adalah murah dan mudah.

Bagaimanapun, bagi membangun komputeran DNA, seuntai bebenang DNA harus diujuk dengan betul bagi menghasil penyelesaian yang berkesan. Jika tidak, komputeran DNA berulang kali gagal mengeluarkan hasil yang sama pada permasalahan yang serupa. Reka bentuk jujukan DNA menjadi satu daripada faktor penentu dalam mencapai ketepatan pengiraan dan penyelesaian masalah.

Selain daripada komputeran DNA, reka bentuk jujukan DNA juga diguna dalam kawasan bioteknologi, seperti DNA teknologi nano dan proses penjana jujukan DNA (Reece 2004). Makalah ini mencadangkan Sistem Imun

Buatan (SIB) sebagai asas algoritma reka bentuk jujukan DNA khususnya Algoritma Pemilihan Negatif (APN). APN dilihat mampu menjadi alternatif kepada masalah reka bentuk DNA. Bagaimanapun, algoritma APN sahaja tidak boleh menyelesaikan masalah ini kerana kaedah gelintar dalam ruang keadaan tidak pelbagai. Oleh yang demikian, beberapa modifikasi dilakukan dengan menambah dua mekanisme dari Algoritma Pemilihan Klonal (APK) iaitu pengklonan dan mutasi. Algoritma yang berasaskan dua algoritma SIB ini dipanggil Pemilihan Negatif dan Klonal – untuk jujukan DNA (PNK-D) dianggap sebagai kaedah penyelesaian yang tahan lasak dengan kemampuannya menggelintar ruang keadaan yang lebih luas dengan kepelbagaian susunan penyelesaian.

REKA BENTUK JUJUKAN DNA

Asid deoksiribonukleik (DNA) merupakan jujuk utama yang membentuk bahan genetik (bahan penentu ciri-ciri kebakaan) kebanyakan organisma hidup. DNA dicipta daripada unit mudah yang disusun dalam suatu jujukan dalam molekul yang besar. Susunan atau jujukan unit ini membawa bersamanya maklumat genetik, sama seperti bagaimana susunan aksara di atas muka surat yang membawa maklumat dan makna. Adleman (1994) menunjukkan bagaimana konsep dan sifat DNA dapat diguna sebagai suatu bentuk pengkomputeran dengan menyelesaikan masalah laluan tujuh-titik Hamilton. Oleh itu, kaedah komputeran ini bergantung kepada jujukan DNA yang baik.

Usaha membina kaedah yang berupaya menghasilkan jujukan DNA yang sempurna masih aktif dijalankan. Bidang penyelidikan reka bentuk jujukan DNA terus berkembang dengan penemuan beberapa algoritma dan kaedah untuk mendapatkan jujukan DNA yang baik. Misalnya reka bentuk jujukan dengan mengguna tumpang tindih sub-jujukan diperkenal oleh Seeman (1990) untuk mendapatkan keunikan jujukan DNA. Baum (1999) mencadang suatu kaedah mereka bentuk jujukan DNA yang unik dengan mengelak penggunaan subjujukan yang berbilang, mengekang mengguna neuklotida pada akhir jujukan.

Hartermink et al. (1999) mengguna kaedah carian yang lengkap, disebut *SCAN*, untuk membuat urutan bagi atur cara mutagenesis. Walaupun kaedah ini berjaya menghasilkan

rangkaian DNA yang baik, tetapi memerlukan masa pengiraan yang tinggi. Dalam kaedah ini, fungsi objektif yang diguna iaitu *hairpin*, tenaga bebas (free energy) dan suhuleleh. Kurniawan et al. (2009) mencadangkan kaedah berasaskan himpunan semut atau *Ant Colony Optimization* (ACO) dengan menukar pengoptimuman multi objektif menjadi satu objektif.

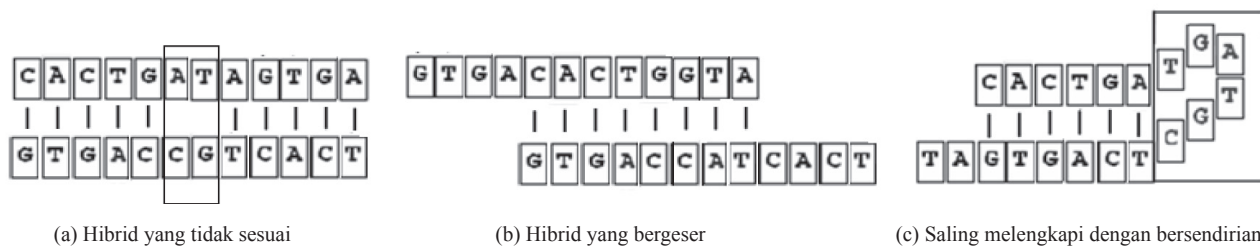
Makalah ini mencadangkan penerapan algoritma Klon dalam permasalahan reka bentuk jujukan DNA, melibatkan algoritma tersebut berinspirasi daripada teori pemilihan klon yang merupakan sebahagian daripada mekanisme sistem imun buatan (SIB). Tahapan proses reka bentuk jujukan DNA disesuai dengan tahapan pada algoritma Klon seperti tahapan tersebut dipapar pada Algoritma 1 dan Rajah 3.

KRITERIA DAN CABARAN REKA BENTUK DNA

Proses penyempurnaan untai DNA adalah penting. Untaian DNA yang stabil adalah untai-pendua yang stabil dalam berhibrid dan tidak saling melengkap antara satu jujukan dengan jujukan yang lain. Terdapat dua kaedah utama dalam menghasil DNA yang stabil: Pertama, Kaedah Pelengkap *Watson-Crick*. Kaedah ini cuba mengurang kadar kebarangkalian terjadinya jujukan DNA yang tidak simetri. Kaedah kedua ialah untuk mengelak daripada terjadinya jujukan DNA hibrid yang tidak sesuai seperti DNA terhibrid dengan cara bersilang-bergeser atau bersilang seperti yang dipapar dalam Rajah 1. Ramai penyelidik tertumpu kepada usaha mengelak daripada terbentuknya DNA yang mempunyai ciri-ciri seperti Rajah 1. Justeru, penyelidikan ini juga tertumpu kepada usaha untuk membangun kaedah mereka bentuk jujukan DNA dengan memerhati kekangan yang dapat mengurang kebarangkalian terjadinya tindak balas yang tidak diinginkan.

PENGIRAAN KEAFINAN

Permasalahan reka bentuk DNA adalah berbeza dengan permasalahan lain yang pernah diselesai dengan mengguna APN mahupun APK kerana kaedah pengiraan keafinan yang berbeza. Konsep keafinan APN dan APK adalah berdasarkan proses perekatan yang berlaku dalam sistem



RAJAH 1. Contoh terjadinya hibrid yang tidak sesuai

imun tabii. Apabila suatu antigen dan reseptor merekat, kadar kekuatan rekatan itu bergantung kepada sejauh mana reseptor kawasan V boleh mengenal pasti antigen. Sebagai gambaran, jika diberi limfosit buatan (x) dan antigen (y), beberapa petua pepadanan boleh digunakan bagi mengukur persamaan antara (x) dan (y).

Kaedah pengiraan keafinan yang diguna dalam permasalahan jujukan DNA adalah berdasarkan empat objektif. Permasalahan pengoptimuman reka bentuk jujukan DNA secara asasnya merupakan kajian yang melibatkan lebih daripada satu objektif. Namun permasalahan multi objektif ini ditukar kepada satu objektif dengan kaedah agregat. Kajian terdahulu merumuskan beberapa objektif yang perlu dipenuhi dalam penghasilan jujukan DNA. Makalah ini mengguna rumus pengukuran objektif yang diguna oleh Kurniawan et al. (2009). Oleh kerana masalah jujukan DNA yang dikaji, makalah ini hanya mengguna satu objektif, maka agregat keempat-empat objektif ini diambil sebagai nilai keafinan. Maka, objektif algoritma cadangan ini adalah untuk mendapat keafinan yang terbaik iaitu keafinan yang rendah nilainya. Ini kerana, masalah penjujukan DNA berusaha menghasil set untaian DNA yang mana setiap untaian adalah unik jujukannya. Maka, terdapat dua pengukuran keafinan: Pertama mengukur keafinan keseluruhan set untaian yang dipanggil keafinan global dan kedua, keafinan lokal iaitu pengukuran keafinan di antara setiap untaian dalam suatu set. Berikut adalah rumus pengukuran objektif yang diguna dalam pengiraan keafinan:

1. $H_{sukatan}$

$H_{sukatan}$ unuk berfungsi mengira jumlah nukleotida yang saling melengkapi dalam usaha untuk mencegah terjadinya hibrid silang antara dua jujukan. Rumusan matematik untuk keserupaan dipapar pada persamaan 1.

$$H_{measure}(x, y) = \max_{|i| < L-1} (H_{dis}(x, shift(rev(y), i)) + H_{con}(x, shift(rev(y), i))). \quad (1)$$

2. Keserupaan

Keserupaan berfungsi mengira nukleotida yang serupa dalam dua untaian DNA yang sejajar, pengiraan tersebut digunakan bagi menjaga keunikan pada setiap untaian DNA yang dijana. Rumusan matematik untuk keserupaan dipaparkan sebagai berikut:

$$f_{similarity}(\Sigma) = \sum_{i=1}^n \sum_{j=1, j \neq i}^n similarity(\Sigma_i, \Sigma_j). \quad (2)$$

3. Kesenambungan

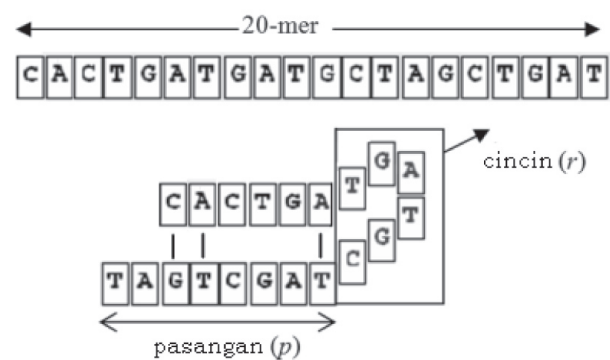
Kiraan jika terdapat pangkal (A, T, C, G) yang terletak dalam jujukan secara berterusan dalam satu untaian DNA. Fungsi kesinambungan untuk satu kumpulan jujukan dapat dirumus seperti berikut (Shin et al. 2005).

$$f_{continuity}(\Sigma) = \sum_{i=1}^n continuity(\Sigma_i). \quad (3)$$

4. Selekoh-tajam

Fungsi selekoh-tajam untuk mengira kebarangkalian terjadinya struktur ganda daripada satu untaian DNA, contoh selekoh-tajam dapat dilihat pada Rajah 2 yang menunjukkan kebarangkalian pembentukan struktur gelang berasaskan pada nilai tertentu dari p (pasangan) dan r (cincin) untuk 20-mer jujukan DNA. Rumusan pengiraan Selekoh-tajam pada persamaan 4.

$$f_{hairpin}(\Sigma) = \sum_{i=1}^n hairpin(\Sigma_i). \quad (4)$$



RAJAH 2. Gambaran bentuk hairpin contoh $p = r = 6$

KEKANGAN DALAM REKA BENTUK DNA

Terdapat dua kekangan yang perlu diperhatikan oleh PNK-D dalam menjana DNA buatan walaupun penjanaan DNA adalah secara rawak. Dua kekangan tersebut adalah suhu leleh dan kandungan GC. Kedua-dua kekangan ini diperhatikan dalam penjanaan DNA untuk memasti DNA yang terhasil adalah stabil dan tidak terurai dengan mudah.

1. Kandungan GC

Kandungan GC merupakan tingkat peratus kandungan G dan C yang terdapat pada untaian DNA. Kandungan G dan C meninggalkan kesan kepada kestabilan sifat kimia untaian DNA. Rumusan kandungan GC ditunjukkan pada persamaan 5. w, x, y, dan z mewakili A, T, G, dan C dalam jujukan per satu.

$$GC_{content} = (yG + zC) / (wA + xT + yG + zG). \quad (5)$$

2. Suhu leleh

Suhu leleh (T_m) juga merupakan faktor penting dalam eksperimen reka bentuk jujukan DNA. Ini menunjukkan sesuatu suhu boleh menyebabkan setengah dari dua untaian terpecah dan terpisah membentuk satu untaian apabila kadar suhu leleh

yang diberi melebihi ambang tertentu. Kajian ini menggunakan model *nearest-neighbour Santa Lucia Unified* sebagai rumusan dalam mengira suhu leleh tersebut. Rumusan kandungan suhu leleh ditunjukkan pada persamaan 6.

$$T_m(X) = \frac{\Delta H}{\Delta S + R \ln C_T} + 16.6 \log(Na^+) \quad (6)$$

Dengan ΔH dan ΔS merupakan perubahan *enthalpy* dan *entropy* daripada tindak balas sepuh lindap seperti yang ditunjukkan pada Jadual 1, C_T ialah tumpuan keseluruhan *oligonucleotides* untaian. Untuk molekul penyempurna-bukan-diri C_T digantikan dengan $C_T/4$. Na^+ ialah tumpuan garam. Seluruh gas-constant - constant (Boltzmann's constant), R , ialah $\left(\frac{1.987 \text{ cal}}{\text{mol}^\circ\text{C}}\right)$.

KAEDAH CADANGAN BERASASKAN SISTEM IMUN BUATAN

Sistem imun buatan (SIB) ditakrif sebagai paradigma pengkomputeran yang diinspirasi daripada fungsi, prinsip dan mekanisme sistem imun tabii (Leandro 2002). Sistem imun tabii mempunyai beberapa mekanisme asas. Setiap mekanisme tersebut mempunyai prinsip kerja yang berbeza seperti pemilihan klonal, pemilihan negatif dan teori merbahaya. Kaedah cadangan yang dibincang dalam makalah ini adalah berasaskan teori pemilihan negatif dan klonal.

Titik tumpuan para penyelidik dalam kebanyakan penyelidikan berasaskan SIB adalah bagaimana limfosit (sel-B dan set-T) menjadi matang, menyesuaikan diri, bertindak dan belajar ketika bertindak balas terhadap antigen asing. Algoritma berasaskan-populasi seperti algoritma pemilihan negatif (APN) dan algoritma pemilihan klonal (APK) tertumpu kepada penjanaan populasi limfosit, memperbaiki dan menapis populasi dengan mengguna teknik yang dipelagak daripada sistem imun tabii. Setelah sel-T (sel timus) terhasil, sel-T berhijrah ke timus untuk dimatangkan dengan mempelajari konsep 'diri.' Proses pembelajaran berlaku melalui dua proses pemilihan iaitu pemilihan positif dan pemilihan negatif.

ALGORITMA PEMILIHAN NEGATIF ASAL

Forrest et al. (1994) membangunkan APN yang diinspirasi oleh teori pemilihan negatif. Sel T limfosit matang dalam timus sebelum dilepas ke dalam sistem peredaran darah. Sel T limfosit yang dimasukkan ke dalam sistem peredaran darah mempunyai ciri-ciri yang menarik iaitu sel T ini hanya boleh diaktif oleh sel asing. Komponen penting APN adalah ingatan yang mengandungi tiga set iaitu set P, M dan C. Set P ialah set-diri dan matlamat APN ialah untuk menyediakan suatu set pola P yang perlu dilindungi. Set M atau set pengesan (M) bertanggungjawab untuk mengenal pasti semua elemen yang menghasilkan set yang mengandungi elemen bukan-diri.

APN bermula dengan menjana calon elemen (C) secara rawak. Setiap elemen dalam C dibandingkan dengan elemen dalam P. Jika wujud persamaan, iaitu apabila elemen P dikenal pasti oleh elemen dalam C, maka elemen C berkenaan dibuang. Jika elemen C tidak mengenal pasti mana-mana elemen P, maka simpan elemen ini dalam set pengesan M. Setelah menjana set pengesan M, langkah seterusnya di dalam APN adalah untuk memantau sistem kehadiran sel bukan-diri atau antigen. Set P mengandungi sel yang perlu dilindungi, set P ini boleh terbentuk dari set P dan sel baharu, atau set semua ahli set P adalah baharu. Set P dan Set M diguna bagi mengenal pasti sel-asing di dalam sistem.

PEMILIHAN NEGATIF DAN KLONAL UNTUK MEREKA BENTUK DNA (PNK-D)

Berasaskan pada APN dan APK, satu algoritma baharu dihasilkan bagi mereka bentuk jujukan DNA. Algoritma ini dipanggil Algoritma berasaskan Pemilihan Klonal dan Negatif untuk mereka bentuk DNA atau PNK-D. Perbezaan utama di antara PNK-D dan APN adalah PNK-D direka bagi menghasilkan pola P yang tidak diguna bagi mengecam sel asing. Set P PNK-D dijana dengan memastikan setiap elemen atau untaian set P (timus) mempunyai keafinan global dan keafinan lokal baik (yakni rendah) untuk diguna oleh komputeran DNA. Namun APN sahaja dilihat tidak mampu untuk menghasilkan kanjuruhan DNA yang baik kerana kaedah penghasilan limfosit APK adalah secara rawak sahaja. Oleh yang demikian, penyelidik meminjam dua mekanisme APK iaitu pengklonan dan mutasi.

Teori Pemilihan Klonal dicadang oleh Burnet (1959). Teori ini adalah berasaskan gerak balas sistem imun kepada rangsangan atau serangan antigen. Teori ini berpendapat bahawa hanya sel yang mampu mengenali suatu antigen berkembang dari segi jumlah. Berasaskan teori Pemilihan Klonal, algoritma *Clonalg* dicadangkan oleh de Castro dan Von Zuben (2002) bagi aplikasi pada bidang pembelajaran mesin dan pengoptimuman. Matlamat membenam dua mekanisme adalah bagi mempelbagai halaaan gelintaran. Penghasilan jujukan DNA yang berkeafinan rendah adalah masalah pengoptimuman yang melibatkan gelintaran dalam ruang keadaan/penyelesaian. Pengklonan membolehkan PNK-D mempelbagai titik gelintaran pada seluruh ruang penyelesaian. Pemilihan titik gelintaran ini pula dibuat secara rawak. Proses mutasi pula adalah satu proses eksploitasi bagi mengelak gelintaran terperangkap dalam keadaan yang dipanggil *local optima*.

PENERAPAN ALGORITMA PNK-D PADA REKA BENTUK JUJUKAN DNA

Dalam penerapannya, proses reka bentuk jujukan DNA dilaksanakan mengikut tahapan algoritma PNK-D seperti yang dipapar pada Algoritma 1 dan Rajah 3. Bermula dari penjanaan jujukan secara rawak pada set P. Kemudian jujukan ada pada set P disalin ke dalam set C. Dalam set C pengiraan keafinan sebagai *local optima* dilaksana

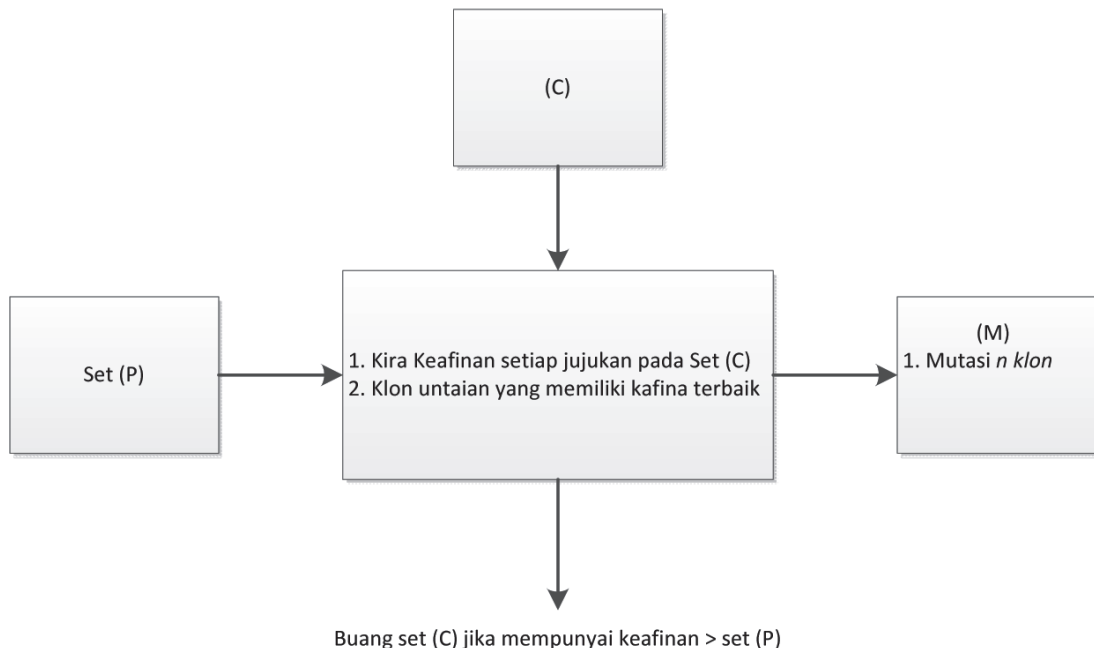
sehingga untaian terbaik diperoleh. Untai DNA terbaik ini ditandai dengan kadar nilai keafinannya. Pada setiap lelaran, untaian terbaik dipilih untuk diganda (klon) dan kemudiannya dimutasi sebagai suatu populasi calon penyelesaian. Untaian yang terbaik pada setiap iterasi mengganti untaian yang paling tidak baik dalam set C. Setelah itu, keafinan global dikira. Set C yang mempunyai keafinan global terendah (terbaik) dipindah ke dalam Set P. Set P dikemas kini dengan membuang untaian set P yang terdahulu. Proses ini diinspirasi daripada kaedah Clonalg yang dicadang oleh de Castro dan Von Zuben (2002). Clonalg menilai antibodi baharu

yang dihasil dan antibodi terbaik ditambah ke dalam populasi asli. Setelah itu, beberapa antibodi yang mempunyai nilai paling teruk diganti dengan antibodi yang berasal daripada populasi baharu yang mempunyai nilai lebih baik (De Castro & Von Zuben 2002). Perlu diingatkan bahawa set P dianalogi sebagai timus yang mengandungi limfosit yang perlu dimatangkan.

PNK-D menerus proses di atas dalam beberapa lelaran yang ditetapkan dengan matlamat menyediakan suatu set jujukan DNA yang boleh diguna oleh komputer DNA yang dipanggil set-diri (P). Algoritma PNK-D adalah seperti berikut:

Algoritma 1: PNK-D untuk reka bentuk jujukan DNA

Tetapkan parameter iterasi r , n dan p
 Jana populasi P secara rawak sebanyak p untaian dengan mematuhi kekangan
 Ulang selagi iterasi $< r$
 Salin populasi set P ke dalam set C
 Kira keafinan lokal setiap elemen C
 Elemen C yang terbaik diklon sebanyak n untaian dengan mematuhi kekangan
 Semua n klon masukkan ke dalam M.
 Mutasi setiap n untaian dalam M
 Untuk setiap n untaian dalam M
 Ganti elemen C yang terendah keafinan lokalnya
 Kira keafinan global C
 Kemaskini set untaian P dengan set untaian C yang mempunyai keafinan global yang tertinggi.
 Kosongkan C dan M
 $r++$



RAJAH 3. Carta Alir Algoritma PNK-D

EKSPERIMEN

Ekperimen dilakukan dengan pendekatan PNK-D yang dicadangkan. Nilai kekangan kandungan GC berkisar antara 20%-80% dan nilai ambang suhu leleh yang diberi adalah di antara 30°C-80°C. Pengiraan suhu leleh adalah berasaskan kaedah *nearest-neighbor* (NN) dengan 1M tumpuan garam dan 10 nM tumpuan jujukan dan panjang DNA yang diguna adalah 20-mer dengan jumlah populasi p adalah 7 untaian. Ringkasan nilai parameter algoritma PNK-D dan parameter daripada jujukan DNA dipapar dalam Jadual 1 dan Jadual 2.

Dalam kajian ini, eksperimen dilarikan dengan suatu simulasi. Nilai keafinan yang diukur adalah hasil agregat nilai objektif yang diambil dari empat objektif iaitu $H_{sukatan}$ keserupaan, kesinambungan dan selekoh-tajam. Simulasi dilarikan pada komputer dengan pemprosesan 1.86 GHz dan 2GB RAM. Hasil yang diperoleh dipapar dalam Jadual 3. Untuk melakukan perbandingan secara adil, maka parameter DNA yang diguna oleh PNK-D adalah sama dengan parameter yang diguna oleh (Kurniawan et al. 2009).

JADUAL 1. Parameter algoritma PNK-D

Parameter	Nilai
Penggandaan	30
Jumlah Pemilihan	1 jujukan
Kadar mutasi	20%
Jumlah Jujukan (P_2)	5 jujukan
Lelaran	500

JADUAL 3. Perbandingan capaian AIS dan P-ACO

AIS						
No	Jujukan DNA	Kesinambungan	Selekoh-Tajam	$H_{sukatan}$	Keserupaan	Keseluruhan
1	TTCTCTATTCTTCTTGTCT	0	0	0	58	58
2	TCTGTGTCTCCTTCTCCTA	0	0	4	58	62
3	CCTTCGTTCCCTCCTCGTC	0	0	12	59	71
4	TCTCTCACACTCCTTCTCT	0	0	8	57	65
5	TCCGCTCCCTCCGTCCTTC	9	0	0	60	69
6	ATCTTGTCCTTCTCTCTT	0	0	4	68	72
7	CTTCCATCCTTCACTCTTT	9	0	4	61	74
	Nilai Keseluruhan	18	0	32	421	471
	Purata	2.5	0	4.5	60	67.3
P-ACO						
No	Jujukan DNA	Kesinambungan	Selekoh-Tajam	$H_{sukatan}$	Keserupaan	Keseluruhan
1	GCAGAACACACACCACCAAC	0	0	4	68	72
2	CACACACACACACACAGAA	0	0	10	62	72
3	ACACCACAACACCACATAGC	0	0	10	67	77
4	CAAGAGAACAACAACCAAGC	0	0	9	69	78
5	CCACCACCACCACACTACA	0	0	12	63	75
6	CACACAAGACACCACAACAG	0	0	8	71	79
7	TACAAGACACACAAGACACA	0	0	8	66	74
	Nilai Keseluruhan	0	0	61	466	527
	Purata	0	0	8.71	66	75.3

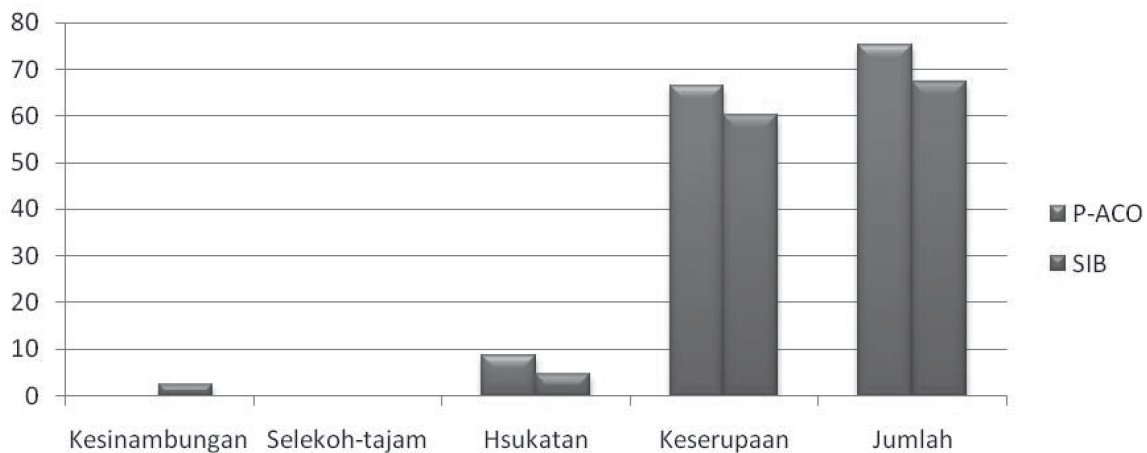
JADUAL 2. Parameter Jujukan DNA

Parameter	Nilai
$H_{sukatan}$	$H_{con} = 6,$ $H_{dis} = 0.17\%$
Keserupaan	$S_{con} = 6,$ $S_{dis} = 0.17\%$
Ambang bersinambung	$t = 2$
Selekoh-tajam	$R_{min} = 6,$ $P_{min} = 6$
Kandungan-GC	20% - 80%
Suhu-leleh (T_m)	30° - 80°
$N_a +$	1M
C_i	10nM

HASIL EKSPERIMEN

Jadual 3 memapar keputusan yang diperoleh dengan nilai purata keseluruhan 67.3. Nilai purata tersebut merupakan nilai purata yang diperoleh dari nilai campuran keseluruhan objektif ($H_{sukatan}$, keserupaan, kesinambungan, selekoh-tajam). Nilai ini dibanding dengan nilai jujukan hasil pemprosesan DNA yang menggunakan kaedah lain seperti kumpulan semut (P-ACO). Perbandingan tersebut dipapar dalam Jadual 3.

Carta perbandingan yang dipapar dalam Rajah 4 memperlihatkan bahawa nilai purata jujukan DNA kaedah AIS lebih kecil daripada nilai purata jujukan DNA yang dijana mengguna kaedah P-ACO. Bagaimanapun, nilai fungsi objektif keserupaan pada jujukan PNK-D adalah lebih tinggi (tidak baik) daripada nilai keserupaan P-ACO.



RAJAH 4. Carta perbandingan hasil dari simulasi

Pemerhatian terhadap perbandingan dalam Rajah 4, mendapati terdapat perbezaan sifat atau karakter pada objektif dan keserupaan. Adalah diperhatikan bahawa ketika nilai rendah maka nilai keserupaan akan tinggi. Sifat daripada kedua-dua objektif tersebut (dan keserupaan) akan berhubung saling terbalik. Hal ini disebabkan oleh perbezaan tujuan yang diguna untuk tugas pengoptimuman, manakala kedua-dua fungsi objektif tersebut mengguna asas perbandingan antara dua jujukan, maka sifat hubungan terbalik tidak dapat dielakkan. Seperti pada contoh, ketika dua jujukan yang terdiri dari "CCAAA", dihibrid dengan "CCAAA" maka nilai dapat dipastikan rendah kerana kedua jujukan tersebut tidak bercantum, tetapi ketika disukat menggunakan fungsi keserupaan, maka peluang untuk terjadinya keserupaan adalah menjadi tinggi.

Kaedah berasaskan-agregat yang diguna dalam kajian ini mempunyai kelebihan dan kelemahan. Kaedah ini merupakan cara menyelesaikan permasalahan pengoptimuman multi objektif dengan menggunakan konsep kebarangkalian secara sederhana. Meskipun demikian, konsep tersebut merupa konsep yang intuitif dan mudah digunakan (Deb 2001).

KESIMPULAN

Algoritma cadangan yang berasaskan dua prinsip penting AIS iaitu APN dan APK untuk permasalahan reka bentuk jujukan DNA dihurai. Penggabungan konsep dan mekanisme APN dan APK adalah bagi memastikan proses pematangan limfosit dalam timus (iaitu set P) dilakukan dengan berkesan kerana pencarian jujukan DNA yang baik dilakukan dalam ruang penyelesaian yang lebih besar. Hasil daripada proses penjanaan DNA buatan mengguna algoritma PNK-D telah dibanding dengan DNA yang diproses menggunakan kaedah pengoptimuman kumpulan semut (ACO). Hasil perbandingan menunjukkan bahawa kaedah yang dicadangkan mempunyai nilai H_{sukatan} dan

keserupaan yang baik, walaupun nilai kesinambungan adalah lebih tinggi berbanding hasil kaedah ACO. Kajian permasalahan pengoptimuman reka bentuk jujukan DNA secara asasnya merupakan kajian yang melibatkan lebih daripada satu objektif. Namun permasalahan multi objektif ini ditukar kepada satu objektif dengan kaedah agregat. Kajian ini membuktikan bahawa PNK-D mempunyai keupayaan yang setara dengan ACO. Justeru, dalam kajian lanjutan, PNK-D perlu diuji dengan penyelesaian masalah multi objektif pula.

RUJUKAN

- Adleman, L. 1994. Molecular Computation of Solutions to Combinatorial Problems. *Science* 266(5187): 1021-1024.
- Burnet, F. 1959. *The Clonal Selection Theory of Acquired Immunity*.: Cambridge University Press.
- De Castro, L. & Von Zuben, F. 2002. Learning and Optimization Using the Clonal Selection Principle. *Evolutionary Computation, IEEE Transactions on* 6(3): 239-251.
- Deb, K. 2001. *Multi-Objective Optimization Using Evolutionary Algorithms*. England: Wiley.
- Forrest, S., Perelson, A. S., Allen, L. & Cherukuri, R. 1994. Self-Nonself Discrimination in a Computer. *Research in Security and Privacy, 1994. Proceedings, 1994 IEEE Computer Society Symposium on*, hlm. 202-212.
- Hartemink, A.J., Gifford, D.K. & Khodor, J. 1999. Automated Constraint-Based Nucleotide Sequence Selection for DNA Computation. *Biosystems* 52(1-3): 227-235.
- Kurniawan, T.B., Ibrahim, Z., Khalid, N.K. & Khalid, M. 2009. *A Population-Based Ant Colony Optimization Approach for DNA Sequence Optimization*. Proceedings of the 2009 Third Asia International Conference on Modelling and Simulation. Anjuran IEEE Computer Society.
- Leandro, C. 2002. *Artificial Immune Systems: A New Computational Intelligence Approach*. Springer-Verlag.
- Reece R. J. 2004. *Analysis of Genes and Genomes*. England: Wiley.
- Seeman, N. 1990. De Novo Design of Sequences for Nucleic Acid Structural Engineering. *Journal of Biomolecular Structure & Dynamics* 8(3): 573-581.

Shin, S.Y., Lee, I.-H., Kim, D. & Zhang, B.-T. 2005. Multiobjective Evolutionary Optimization of DNA Sequences for Reliable DNA Computing. *Evolutionary Computation, IEEE Transactions on* 9(2): 143-158.

M. Daman Huri
Mohd Zakree Ahmad Nazri
Azuraliza Abu Bakar
Universiti Kebangsaan Malaysia
43600 UKM, Bangi
Selangor, Malaysia
damanhuri63@yahoo.co.id
mzan@ftsm.ukm.my
aab@ftsm.ukm.my

Tribasuki Kuniawan
Universitas Bina Darma
Palembang, Sumatera Selatan, Indonesia
tribasukikurniawan@yahoo.com