



Kontribusi Aplikasi Medis dari Ilmu Bioinformatika Berdasarkan Perkembangan Pembelajaran Mesin (*Machine Learning*) Terbaru

Arli Aditya Parikesit

Department of Bioinformatics, School of Life Sciences, Indonesia International Institute for Life Sciences (i3I), Jakarta, Indonesia

ABSTRAK

Berkembangnya ilmu bioinformatika merupakan konsekuensi banyaknya data eksperimen laboratorium para peneliti biologi molekuler ataupun biomedis. Selain pengembangan basis data terpusat yang merupakan kompetensi inti ilmu bioinformatika, pendekatan komputasi lain seperti pembelajaran mesin juga dikembangkan, sehingga data tersebut dapat diolah menjadi informasi yang berguna bagi dunia kesehatan. Kajian ini akan menelaah perkembangan pendekatan pembelajaran mesin pada ilmu bioinformatika, dan aplikasinya pada dunia kesehatan terutama pada informatika kanker dan virus. Masa depan aplikasi medis dengan ilmu bioinformatika menarik karena melibatkan berbagai pendekatan baru seperti kecerdasan buatan dan biologi sistem.

Kata kunci: Bioinformatika, biomedis, biologi molekuler, biologi sistem, kecerdasan buatan, pembelajaran mesin

ABSTRACT

The development of bioinformatics science is a consequence of the massive data generation of laboratory experiments conducted by molecular biology and biomedical researchers. In addition to the development of a centralized database that is the core competence of bioinformatics science, other computing approaches such as machine learning are also developed so that the data can be processed into useful information for the human health. This review will examine the development of machine learning approaches in bioinformatics science, and its application to the human health, especially in cancer and virus informatics. The future of medical applications with bioinformatics science is exciting as it involves various new approaches such as artificial intelligence and system biology. **Arli Aditya Parikesit. The Contribution of Medical Application from Bioinformatics Based on Latest Advances of Machine Learning**

Keywords: Artificial intelligence, biomedical, bioinformatics, machine learning, molecular biology, system biology

Bioinformatika dan Pembelajaran Mesin

Perkembangan ilmu kedokteran modern tidak dapat dipisahkan dari perkembangan ilmu biologi molekuler. Sejak Watson dan Crick menemukan struktur DNA, perkembangan ilmu biologi molekuler menjadi sangat pesat. Terjadi lonjakan jumlah data dari eksperimen laboratorium pada laboratorium biologi molekuler, terutama pada laboratorium penelitian biomedis atau kedokteran dasar.¹ Data terutama berupa sekuens atau urutan asam nukleat dan protein.² Data urutan biologis tersebut perlu diolah menjadi informasi yang bermanfaat untuk pengembangan agen profilaksis dan terapeutik penyakit. Sebelum ditemukannya *Personal Computer* atau PC, sumber daya komputasi sangat terbatas, sehingga pengolahan data biologis dilakukan secara

manual.³ Dengan mulai dipopulerkannya IBM PC dan kompatibelnya pada awal tahun 1980-an, peneliti biomedis mulai mendapatkan akses terhadap sumber daya komputasi yang memadai untuk mengolah urutan biologis. Dari perkembangan tersebut, lahir ilmu bioinformatika.⁴

Bioinformatika didefinisikan sebagai ilmu gabungan antara biologi molekuler dan teknik informatika.⁵ Tujuan pengembangan ilmu baru tersebut adalah untuk mempermudah pengolahan data urutan biologis dengan metode komputasi terbaru, sehingga mendapatkan informasi atau anotasi signifikan untuk pengembangan ilmu kedokteran ataupun cabang ilmu hayati lainnya.⁶

Ilmu bioinformatika mencapai titik awal

kemajuan yang sangat signifikan sejak Proyek Genom Manusia (*Human Genome Project/HGP*) diluncurkan pada awal tahun 1990-an.⁷ Proyek ambisius ini bertujuan mengurutkan asam nukleat dan protein dari genom dan proteome manusia, sehingga dapat membuka kemungkinan baru untuk anotasi informasi bagi perkembangan ilmu biosains/hayati.⁸ Namun, selesainya HGP pada awal 2000-an tidaklah menyelesaikan permasalahan kedokteran begitu saja, justru membuka pertanyaan-pertanyaan baru. HGP menunjukkan bahwa mekanisme molekuler pada sel tidaklah cukup dipelajari melalui kajian genomik dan proteomik, namun juga harus lebih mendalami kajian baru lain seperti transkriptomik dan epigenomik.⁹ Bersamaan dengan dimulainya HGP, lembaga penelitian di Amerika Serikat (NCBI), Jepang (DDBJ), dan

Alamat Korespondensi email: arli.parikesit@i3l.ac.id



Uni Eropa (EBI) mengembangkan basis data urutan/sekuens biologis yang bermanfaat untuk menyimpan, mengolah, menganalisis, dan menyajikan informasi yang dapat dimanfaatkan untuk aplikasi biologi molekuler pada berbagai cabang ilmu hayati, seperti kedokteran, lingkungan, ataupun pertanian.¹⁰ Besarnya jumlah data yang disimpan dan diolah pada basis data tersebut menyebabkan diperlukannya pendekatan baru dalam mengelola 'ledakan informasi' tersebut, terutama di era informasi seperti sekarang ini.

Perkembangan pesat teknik informatika (TI) dengan adanya internet, basis data di awan (*cloud*), dan gawai cerdas (*smart gadget*) mengharuskan pengelolaan data yang sangat besar atau '*big data*' dikelola dengan pendekatan dan metode yang tepat juga. Salah satu metode yang dikembangkan adalah pendekatan pembelajaran mesin atau *machine learning*, yang memungkinkan prediksi fungsi dan struktur asam nukleat ataupun protein dengan resolusi dan ketepatan sangat baik.¹¹ Beberapa metode pembelajaran mesin yang umum digunakan adalah *hidden markov model* (HMM), *artificial neural network* (ANN), dan *support vector machine* (SVM), metode yang paling banyak digunakan adalah ANN.¹² Pada dasarnya, pendekatan pembelajaran mesin bekerja dengan aplikasi algoritma yang mendorong komputer untuk melakukan operasi berdasarkan data empiris yang dilatih pada basis data terintegrasi. Pola atau profil yang dikembangkan dari pelatihan tersebut akan digunakan untuk prediksi pola asam nukleat atau protein tertentu yang masih baru. Pada titik ini, pembelajaran mesin dapat digunakan untuk mendeteksi pola prevalensi penyakit infeksi ataupun kanker pada sampel, misalnya pada metode instrumentasi *DNA Microarray*. Selain untuk keperluan diagnostik, pembelajaran mesin juga dapat digunakan untuk menemukan kandidat obat dan vaksin yang mungkin dikembangkan pada eksperimen basah. Basis data untuk keperluan tersebut sudah tersedia di dunia maya, bahkan sudah diaplikasikan untuk penelitian bidang kesehatan.^{13,14}

Informatika Virologi

Infeksi virus merupakan salah satu manifestasi penyakit yang mempengaruhi dunia kedokteran dengan sangat signifikan. Sebagai parasit intraseluler obligat, virus dapat didefinisikan sebagai entitas yang tidak hidup,

namun juga tidak mati. Virus mempunyai kapabilitas membajak sistem informasi genetik sel, sehingga mampu melakukan modifikasi sistem metabolisme sel tersebut. Contoh paling terkenal adalah virus HIV (*human immunodeficiency virus*), penyebab penyakit AIDS (*acquired immunodeficiency syndrome*). Infeksi virus HIV menyebabkan penurunan kapabilitas sistem imun manusia, dengan membajak dan membunuh sel T, sel dendritik, dan makrofag yang berperan dalam sistem tersebut.¹⁵ Untuk mengatasi masalah tersebut, peneliti bioinformatika dan biomedis telah mengembangkan obat HIV/AIDS dengan bantuan metode komputasi termutakhir, yaitu penambatan dan dinamika molekul.¹⁶ Data untuk pengembangan obat tersebut diunduh dari basis data '*HIV-1 Human Interaction Database*' yang dikelola oleh NCBI (<https://www.ncbi.nlm.nih.gov/genome/viruses/retroviruses/hiv-1/interactions/>). Informasi yang tersedia pada basis data tersebut adalah terkait informasi interaksi protein-protein dengan parameter terkait, yaitu: fenotip, data ekspresi gen, kluster homologen, transkrip refseq, dan jalur biologis. Data tersebut bisa dicari dengan melakukan navigasi terhadap gen manusia ataupun HIV.⁷

Informasi data biologis yang didapatkan dari basis data tersebut, terutama informasi urutan protein, akan menjadi titik berangkat bagi prediksi struktur dengan metode *homology modeling*, dimana urutan protein yang ada akan dibandingkan dengan struktur yang ada di *template* basis data. Jika struktur sudah berhasil dielusidasi, maka kandidat obat atau ligan dapat dicari (*search*) pada basis data terkait seperti Pubchem (<https://pubchem.ncbi.nlm.nih.gov>). Metode penambatan dan dinamika molekul akan mencari ligan yang terbaik dalam konteks interaksinya dengan protein. Informasi berbasis pembelajaran mesin ini akan menjadi masukan bagi pengembangan obat atau ligan terbaru di laboratorium. Diperlukan pemahaman yang baik terhadap ilmu biokimia untuk mempelajari interaksi ligan-protein, sebab hal ini melibatkan ikatan kimia, seperti ikatan hidrogen, kovalen, ataupun ionik. Keberadaan ikatan tersebut akan dikaji pada interaksi antara ligan dan residu asam amino.

Fokus penelitian informatika virologi tentu saja tidak hanya kepada HIV/AIDS, namun juga kepada virus lain seperti Dengue, Zika, Ebola,

ataupun lainnya. Informasi keseluruhan virus yang menjadi sumber penyakit berbahaya tersebut disimpan pada basis data '*virus variation source*' milik NCBI (<https://www.ncbi.nlm.nih.gov/genome/viruses/variation/>).

Basis data *virus variation resource* menunjukkan bahwa informasi urutan basa ataupun protein virus telah berhasil dikompilasi pada situs yang terpusat dan komprehensif. Informasi urutan protein virus Ebola telah dapat dimanfaatkan sebagai titik awal pengembangan obat dengan bantuan metode penambatan molekul yang sangat umum digunakan pada kajian bioinformatika.¹⁷ Laju mutasi yang sangat tinggi pada virus berinti RNA, seperti HIV dan Ebola, mendorong ilmuwan untuk mengembangkan pendekatan pembelajaran mesin, seperti metode HMM, untuk memprediksi struktur dan fungsi protein virus tersebut. Salah satu basis data berbasis pembelajaran mesin yang bermanfaat untuk melakukan klasifikasi fungsi dan struktur protein adalah SUPERFAMILY.¹⁸ Basis data ini dikembangkan dengan pendekatan pembelajaran mesin, yang melatih data ratusan urutan protein untuk mengembangkan profil deteksi klasifikasi struktur protein yang *solid* dan terpercaya. Tendensi virus untuk 'membajak' sel inang, dan memproduksi perangkat proteomiknya sendiri menjadikan basis data berbasis pembelajaran mesin penting peranannya dalam mempelajari mekanisme infeksi virus, untuk kemudian memprediksi cara untuk mematahkan siklus hidup mereka. Pendekatan simulasi molekuler telah memungkinkan dilakukan penapisan terhadap pustaka kandidat obat yang besar, sehingga hanya kandidat yang benar-benar terpilih dapat diproses pada eksperimen laboratorium.

Informatika Kanker

Kajian bioinformatika terhadap penyakit kanker pada dasarnya menggunakan data yang tidak jauh berbeda dengan virus, yaitu anotasi biomolekul (DNA, RNA, dan protein). Hanya saja, karena ekspresi genetik dan metabolik kanker yang tentu saja umumnya berbeda dengan infeksi virus, data tersebut diolah dengan pendekatan yang tidak sepenuhnya sama. Ada kasus infeksi virus dan kanker mengalami persilangan dan korelasi erat, seperti pada kanker serviks yang disebabkan oleh virus HPV (*Human Papilloma Virus*). Metode simulasi molekuler pengembangan



obat penyakit tersebut dilakukan dengan pendekatan yang tidak jauh berbeda dengan infeksi virus lainnya,¹⁹ bahkan vaksin HPV/kanker serviks juga sudah dikembangkan dan dijual.²⁰ Namun, penyebab penyakit kanker tidak hanya karena infeksi virus, akan tetapi juga dapat oleh faktor keturunan. Basis data OMIM (<http://omim.org>) menyimpan informasi mengenai korelasi antara penyakit kanker dan faktor keturunan (genetik).

Selain itu, penyakit kanker tentu saja dapat disebabkan oleh mutasi genetik atau perubahan urutan basa pada gen, yang informasinya disimpan pada basis data ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>). Mutasi genetik pada penyakit kanker dapat disebabkan oleh faktor keturunan, infeksi virus, ataupun pencemaran lingkungan.²¹ Sangat berbeda dengan infeksi virus, informasi dari basis data OMIM ataupun ClinVar seyogyanya menjadi pertimbangan dalam pengembangan instrument baik diagnostik maupun terapeutik bagi penyakit kanker. Dengan demikian, penyakit kanker dapat ditangani dengan lebih baik.

Salah satu obat kanker yang dibantu pengembangannya dengan ilmu bioinformatika adalah nimotuzumab, yang digunakan sebagai agen kemoterapi untuk terutama menghadapi kanker berbasis *squamous cell carcinoma* atau yang berasal dari sel *squamus*. Sebagai antibodi monoklonal, interaksi nimotuzumab dengan *epidermal growth factor receptor* (EGFR) dipelajari dengan *tool* bioinformatika, untuk kemudian menjadi basis informasi untuk pengembangan di *wet laboratory*.²² Diharapkan, pendekatan multidisiplin antara ilmu bioinformatika dan penelitian eksperimental laboratorium akan makin banyak, sehingga akan memiliki dampak signifikan pada perkembangan ilmu pengetahuan, termasuk aplikasinya.

Aspek pembelajaran mesin pada informatika kanker terutama dapat dilihat pada konteks pengembangan obat; dapat dikembangkan basis data interaksi protein-obat yang lengkap. Profil matriks dapat diciptakan dari interaksi tersebut, sehingga dapat digunakan untuk memprediksi pengembangan obat yang sama sekali baru.²³ Hanya saja, karena penyakit kanker sangat kompleks, aspek jalur metabolisme harus diperhatikan dengan

baik.²⁴ Menghambat sebuah protein yang bertanggung jawab pada jalur metabolik penyakit kanker belum tentu menghentikan proliferasi kanker, karena jalur metabolik lain akan dilibatkan juga.²⁵ Di sini, pengembangan basis data metabolisme juga diperlukan untuk memperkuat informasi bagi pengembangan obat, seperti dicontohkan oleh proyek KEGG.²⁶ Basis data tersebut memiliki kepustakaan jalur metabolik penyakit yang sangat lengkap, bahkan termasuk informasi mengenai metabolisme *drug resistant* dan interaksi antara obat dan protein. Perangkat yang tersedia pada basis data KEGG dapat menjadi titik awal untuk melakukan simulasi molekuler yang lebih canggih, seperti melakukan penambatan dan dinamika molekuler.

Masa Depan Bioinformatika Medis

Berbagai faktor dinamika ekologis ataupun lingkungan di dunia menyebabkan lahirnya penyakit baru. Kemajuan kajian bioteknologi dan biologi molekuler belum dapat sepenuhnya menangani permasalahan penyakit baru tersebut, seperti pada fenomena proses penuaan (*aging*), kanker, HIV/AIDS, dan sebagainya. Bahkan penyakit lama seperti TBC menjadi masalah lagi karena kuman tersebut menjadi resisten terhadap antibiotik. *Wet laboratory* telah menghasilkan banyak data eksperimen biomedis, namun nampaknya solusi komprehensif pengolahan data untuk kepentingan kesehatan masih harus dioptimalkan. Pendekatan biologi yaitu interaksi dan hierarki sistem biologi dipelajari dari tingkat atom sampai ke tingkat ekosistem, merupakan salah satu pendekatan yang menjanjikan untuk diterapkan dalam dunia kesehatan. Penerapan '*big data*' sangat diperlukan, karena biologi sistem mengolah data dari semua aspek yang memungkinkan, bahkan dari berbagai sudut pandang keilmuan, untuk menghasilkan informasi yang dapat dimanfaatkan untuk dunia kesehatan. Hal ini penting, sebab prevalensi penyakit merupakan *event* yang sangat multi-faktorial, sehingga diperlukan pendekatan yang dapat merangkum semua indikator dan parameter ke dalam suatu sistem informasi komprehensif. Ilmu bioinformatika dapat berperan penting karena memiliki kompetensi memadai untuk mengolah informasi tersebut.

Sampai sekarang, pusat informasi basis data genom masih didominasi oleh Amerika

Serikat (NCBI), Uni Eropa (EBI), dan Jepang (DDBJ),^{3,8} Indonesia mulai mengembangkan basis data genom, seperti yang dilakukan oleh litbang pertanian (<http://genom.litbang.pertanian.go.id>) Peneliti kita juga sudah mengembangkan basis data obat bahan alam ataupun jamu (<http://herbaldb.farmasi.ui.ac.id/v3/> dan <http://kanaya.naist.jp/jamu/top.jsp>); yang terakhir bekerja sama dengan Jepang. Maraknya pengembangan basis data genom dan bahan alam merupakan awal yang sangat bagus bagi akselerasi ilmu bioinformatika di Indonesia. Selanjutnya diharapkan bisa dikembangkan basis data yang lebih tersentralisir, seperti NCBI di Amerika Serikat, yang menjadi pusat rujukan nasional.

Di era industri 4.0 ini, dunia kesehatan telah bergerak menuju era *artificial intelligence* (AI/kecerdasan buatan). AI memungkinkan pengembangan metode prediksi pencegahan, terapi, ataupun prevensi penyakit dengan lebih cerdas, karena tidak perlu pelatihan data dalam jumlah besar. Pendekatan AI juga telah diaplikasikan untuk mendesain antibiotik peptida untuk menghadapi '*superbugs*' dengan aktivitas biologis yang diprediksi secara tepat oleh metode komputasi.²⁷ Diharapkan, akan makin banyak aplikasi AI untuk bidang kesehatan, terutama untuk memprediksi luaran dari eksperimen laboratorium. Di sisi lain, ilmu teknik informatika mengalami kemajuan pesat dalam kemampuan simulasi molekuler semenjak *Graphics Processing Unit* (GPU) diperkenalkan secara luas. Sebagai unit pendukung pada komputer, GPU memberikan kinerja yang dalam hal tertentu mendekati komputer super, sehingga dapat melakukan simulasi molekuler dalam ukuran besar, baik untuk penjejakan urutan maupun struktur biomolekul.²⁸ Diharapkan, teknologi GPU dapat membantu pengembangan alat diagnostik, obat, dan vaksin dengan mengakselerasi kinerja prediksi perangkat lunak yang tersedia.

Ucapan Terima Kasih

Terima kasih kepada LPPM ISI atas dukungannya terhadap penulisan manuskrip ini.



DAFTAR PUSTAKA

1. Brusil V. The growth of bioinformatics. Briefings in bioinformatics [Internet]. 2007 Dec 19 [cited 2016 Sep 6];8(2):69–70. Available from: <http://bib.oxfordjournals.org/cgi/doi/10.1093/bib/bbm008>
2. Hagen JB. The origins of bioinformatics. *Nat Rev Genet* [Internet]. 2000 Dec [cited 2015 Jan 28];1(3):231–6. Available from: <http://dx.doi.org/10.1038/35042090>
3. Claverie JM, Notredame C. Bioinformatics for dummies [Internet]. 2011. Available from: <http://books.google.co.id/books?id=4Tw0aZBnBLEC>
4. Tambunan USF, Parikesit AA. Cracking the genetic code of human virus by using open source bioinformatics tools. *Malaysian J Fundam Appl Sci* [Internet]. 2010;6(1):42–50. Available from: <http://mjfas.ibnusina.utm.my/index.php/jfs/article/view/41>
5. Welch L, Lewitter F, Schwartz R, Brooksbank C, Radivojac P, Gaeta B, et al. Bioinformatics curriculum guidelines: Toward a definition of core competencies. *PLoS Comput Biol* [Internet]. 2014 Mar 6 [cited 2016 Nov 28];10(3):e1003496. Available from: <http://dx.plos.org/10.1371/journal.pcbi.1003496>
6. Serapião CNF, AB de S. Bioinformatics: Strategies, trends, and perspectives. In: Lazinica A, editor. *New advanced technologies* [Internet]. 2010 [cited 2014 Aug 13]:126–14-. Available from: <http://www.intechopen.com/books/new-advanced-technologies/bioinformatics-strategies-trends-and-perspectives>
7. Collins FS, McKusick VA. Implications of the human genome project for medical science. *JAMA* [Internet]. 2001 Feb 7 [cited 2013 Aug 16];285(5):540–4. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11176855>
8. Parikesit AA, Anugoro D, Putranto RA. Pemanfaatan bioinformatika dalam bidang pertanian dan kesehatan (The utilization of bioinformatics in the field of agriculture and health). *E-Journal Menara Perkeb* [Internet]. 2017 Oct 30 [cited 2017 Oct 31];85(2). Available from: <http://mp.iribb.org/index.php/mpjurnal/article/view/237>
9. Lunshof JE, Bobe J, Aach J, Angrist M, Thakuria JV, Vorhaus DB, et al. Personal genomes in progress: From the human genome project to the personal genome project. *Dialogues Clin Neurosci* [Internet]. 2010 Jan [cited 2013 Aug 16];12(1):47–60. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3181947&tool=pmcentrez&rendertype=abstract>
10. NCBI. The NCBI handbook [Internet]. 2013 [cited 2013 Aug 1]. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK21101/>
11. Kumar D, Sarvate A, Singh S, Priya P. Virtual screening using machine learning approach. In: 2013 IEEE conference on information and communication technologies [Internet]. 2013 [cited 2014 Mar 10]:594–9. Available from: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6558164>
12. Juhl Jensen L, Bateman A. The rise and fall of supervised machine learning techniques. *Bioinformatics* [Internet]. 2011;27:3331–2. Available from: <http://bioinformatics.oxfordjournals.org/cgi/doi/10.1093/bioinformatics/btr585>
13. Wishart DS, Knox C, Guo AC, Cheng D, Shrivastava S, Tzur D, et al. DrugBank: A knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res*. 2008;36(Database issue):901–6.
14. Moise L, De Groot AS. Putting immunoinformatics to the test. *Nat Biotechnol* [Internet]. 2006 Jul [cited 2015 Jan 28];24(7):791–2. Available from: <http://dx.doi.org/10.1038/nbt0706-791>
15. WHO. World Health Organization HIV/AIDS resource center [Internet]. 2018 [cited 2018 Apr 26]. Available from: <http://www.who.int/hiv/en/>
16. Kuroda DG, Bauman JD, Challa JR, Patel D, Troxler T, Das K, et al. Snapshot of the equilibrium dynamics of a drug bound to HIV-1 reverse transcriptase. *Nat Chem* [Internet]. 2013 Mar [cited 2018 Apr 11];5(3):174–81. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23422558>
17. Tambunan USF, Alkaff AH, Nasution MAF. Bioinformatics approach to screening and developing drug against Ebola. In: *Advances in Ebola control* [Internet]. 2018 [cited 2018 May 2]:75–88. Available from: <https://www.intechopen.com/books/advances-in-ebola-control/bioinformatics-approach-to-screening-and-developing-drug-against-ebola>
18. De Lima Morais DA, Fang H, Rackham OJL, Wilson D, Pethica R, Chothia C, et al. SUPERFAMILY 1.75 including a domain-centric gene ontology method. *Nucleic Acids Res* [Internet]. 2011 Jan 9 [cited 2014 Jul 10];39(Database issue):427–34. Available from: <http://nar.oxfordjournals.org/content/early/2010/11/08/nar.gkq1130.short>
19. Tambunan USF, Bakri R, Parikesit AA, Ariyani T, Puspitasari RD, Kerami D. In silico modification of Zn²⁺ binding group of suberythranilide hydroxamic acid (SAHA) by organoselenium compounds as Homo sapiens class II HDAC inhibitor of cervical cancer. *IOP Conf Ser Mater Sci Eng* [Internet]. 2016 Feb 5 [cited 2016 Mar 17];107(1):012054. Available from: <http://iopscience.iop.org/article/10.1088/1757-899X/107/1/012054>
20. Michels KB, zur Hausen H. HPV vaccine for all. *Lancet* [Internet]. 2009 Jul 25 [cited 2013 Nov 20];374(9686):268–70. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19586657>
21. American Cancer Society. *Cancer facts & figures 2013*. Vol. 2010, American Cancer Society. 2013.
22. Tundidor Y, García-Hernández CP, Pupo A, Infante YC, Rojas G. Delineating the functional map of the interaction between nimotuzumab and the epidermal growth factor receptor. *MAbs* [Internet]. 2014 Jul 23 [cited 2018 May 4];6(4):1013–25. Available from: <http://www.tandfonline.com/doi/abs/10.4161/mabs.28915>
23. Whittaker PA. What is the relevance of bioinformatics to pharmacology? Vol. 24, *Trends in Pharmacological Sciences*; 2003. p. 434–9.
24. Desany B, Zhang Z. Bioinformatics and cancer target discovery. Vol. 9, *Drug Discovery Today*; 2004. p. 795–802.
25. Lin JH, Lu AY. Role of pharmacokinetics and metabolism in drug discovery and development. *Pharmacol Rev* [Internet]. 1997 Dec [cited 2013 Apr 15];49(4):403–49. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9443165>
26. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* [Internet]. 2017 Jan 4 [cited 2018 May 4];45(D1):D353–61. Available from: <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkw1092>
27. Cherkasov A, Hilpert K, Jenssen H, Fjell CD, Waldbrook M, Mullaly SC, et al. Use of artificial intelligence in the design of small peptide antibiotics effective against a broad spectrum of highly antibiotic-resistant superbugs. *ACS Chem Biol* [Internet]. 2009 Jan 16 [cited 2018 May 4];4(1):65–74. Available from: <http://pubs.acs.org/doi/abs/10.1021/cb800240j>
28. Suhartanto H, Yanuar A, Wibisono A. Performance analysis cluster and GPU computing environment on molecular dynamic simulation of BRV-1 and REM2 with GROMACS. 2012 Oct 16 [cited 2013 Aug 16]; Available from: <http://arxiv.org/abs/1210.4251>