p-ISSN: 1693-2617 e-ISSN: 2528-7613

Molecular Characterization On Vitellogenins (VTG) of Grass Carp (Ctenopharyngodon Idella)

Silmy Aulia Rufiatin Nisa^{1)*}, Nur Apriatun Nafisah²⁾, Christosie Immanuel Wahyudi³⁾

^{1)*} Universitas Jenderal Soedirman, Fakultas Biologi, Purwokerto, Indonesia, silmy.nisa@unsoed.ac.id

²⁾ Universitas Jenderal Soedirman, Fakultas Biologi, Purwokerto, Indonesia, nur.nafisah@unsoed.ac.id

³⁾ Universitas Jenderal Soedirman, Fakultas Biologi, Purwokerto, Indonesia, christosie.wahyudi@unsoed.ac.id

Abstrak

Ikan koan memiliki berbagai manfaat untuk kepentingan akuakultur karena mampu berperan sebagai agen biokontrol di perairan sekaligus sebagai sumber protein berkualitas tinggi sehingga berpotensi sebagai target pengembangan budidaya ikan. Upaya yang dapat dilakukan dalam mengembangkan budidaya perikanan adalah menghasilkan bibit unggul melalui manipulasi reproduksi. Penelitian ini dilakukan untuk memahami karakteristik molekuler serta biokimia dari protein Vtg sebagai salah satu protein yang berperan penting dalam oogenesis pada ikan koan, khusunya selama proses Vitelogenesis. Studi dilakukan secara in silico, antara lain analisis filogenetik pada ketiga jenis protein Vtg, analisis karakter fisikokimia dan prediksi domain serta situs fosforilasi. Hasil karakterisasi menunjukkan bahwa Ci-VtgAa dan Ci-VtgAb memiliki hubungan evolusioner yang lebih dekat dibanding dengan Ci-VtqC. Hal tersebut didukung dengan analisis prediksi domain protein yolk, dimana Ci-VtqC tidak memiliki domain Phosvitin (Pv). Hasil analisis karakter fisikokimia mengindikasikan bahwa protein Ci-VtgC cenderung bersifat asam namun lebih stabil dibandingkan Ci-VtgAa dan Ci-VtgAb. Ketiga protein memiliki ketahanan terhadap rentangan temperatur yang luas, memiliki interaksi yang baik dengan molekul air, serta komposisi asam amino didominasi oleh Leucine (Leu). Adanya situs fosforilasi dan O-glikosilasi pada ketiga protein mengindikasikan peran Ci-VtgAa, Ci-VtgAb, dan Ci-VtgC dalam transpor ion logam dan karbohidrat.

Kata Kunci: ikan koan, betina, analisis protein, vitelogenesis

Abstract

Grass carp has various benefits for aquaculture purposes as a biocontrol agent in the waters and also a source of high-quality protein. To develop its aquaculture, we need to produce better fish seeds through manipulation of its reproduction. This study was conducted to understand the molecular and biochemical characteristics of Vtgs as one of the proteins that play an important role in oogenesis in grass carp. The study was conducted in silico, including phylogenetic analysis on the three types of Vtg proteins, physicochemical character analysis, and prediction of domains and phosphorylation sites. Characterization results showed that Ci-VtgAa and Ci-VtgAb have a closer evolutionary relationship than Ci-VtgC. This data was supported by the domain prediction analysis of yolk proteins, where Ci-VtgC does not have a Phosvitin (Pv) domain. The results of physicochemical character analysis indicated that Ci-VtgC protein tends to be acidic but more stable than Ci-VtgAa and Ci-VtgAb. All three proteins have resistance to a wide temperature range, have good interactions with water molecules, and the amino acid composition is dominated by Leucine (Leu). The presence of phosphorylation and O-glycosylation sites in the three proteins indicates the role of three proteins in metal ion and carbohydrate transport.

Keywords: grass carp, female, protein analysis, vitellogenesis

p-ISSN: 1693-2617 e-ISSN: 2528-7613

PENDAHULUAN

Ikan koan (*Ctenipharyngodon idella*) atau dikenal juga dengan sebutan grass carp merupakan salah satu jenis ikan penting dalam budidaya perikanan air tawar di Indonesia (Pangkey *et al.*, 2013). Ikan ini berasal dari sungai-sungai besar di daerah China, Siberia dan Manchuria yang diintroduksi ke sejumlah negara seperti Malaysia, Thailand, Taiwan, Jepang, Amerika Serikat, dan Eropa Timur (Resmikasari, 2008).

Ikan Koan dikenal sebagai ikan herbivor sehingga dalam pemanfaatannya seringkali digunakan sebagai agen biokontrol terhadap keberadaan gulma sehingga dapat meningkatkan produksi perikanan melalui polikultur. Popularitas penggunaan ikan Koan disebabkan karena kemampuan dan daya tahan tubuh yang cukup tinggi, mudah untuk dibudidayakan, mampu berperan sebagai kontrol biologis yang efektif untuk berbagai tumbuhan air, serta memiliki kelezatan sebagai sumber protein berkualitas tinggi (Cudmore & Mandrak, 2004). Ikan koan memiliki rasa yang enak dengan daging berwarna putih, bertekstur kenyal dan tebal pada seluruh bagian tubuhnya (Choirom, Syahrizal & Safratilofa, 2020).

Adanya budaya masyarakat Indonesia yang mengkonsumsi ikan sebagai bahan makanan menjadikan munculnya tren peningkatan konsumsi ikan setiap tahun (Kementrian Kelautan dan Perikanan, 2018). Oleh karena itu, ketersediaan suplai ikan koan melalui pengembangan budidaya perikanan menjadi hal yang penting. Salah satu upaya yang dapat dilakukan dalam mengembangkan budidaya perikanan adalah menghasilkan benih ikan koan yang berkualitas unggul melalui manipulasi reproduksi. Untuk dapat melakukan manipulasi reproduksi, maka perlu dilakukan langkah awal berupa studi yang lebih mendalam mengenai aspek reproduksi khususnya pada ikan betina, Bagaimanapun, mekanisme vitelogenesis dari ikan koan masih belum dapat dijelaskan secara rinci, sehingga dalam penelitian ini dilakukan penelitian secara in silico untuk mengungkap karakteristik molekuler dari sejumlah protein yang mendukung keberhasilan proses vitelogenesis pada gonad ikan koan betina.

Vitelogenesis merupakan tahapan penting pada oogenesis yang terjadi sebelum pematangan gonad, sehingga Vitelogenesis menjadi bagian dari aspek reproduksi yang menarik untuk dikaji. Pada tahap vitellogenesis, terjadi pengumpulan egg yolk precursor proteins (EYPP), terutama Vitelogenin (Vtg) oleh oosit yang sedang tumbuh untuk diproses menjadi protein yolk di ooplasma. Vtg merupakan glikolipofosfoprotein yang spesifik terdapat pada vertebrata maupun invertebrata betina (200-700 kDa) yang bertelur (Pipil et al., 2015). Pada ikan teleostei, umumnya terdapat tiga jenis dari Vtg yang dibedakan berdasarkan strukturnya, antara lain VtgAa, VtgAb, dan VtgC.

Pada penelitian ini, kami mengkaji Vtg sebagai protein utama yang berperan dalam proses vitelogenesis dengan tujuan untuk memperoleh informasi ilmiah mengenai karakteristik Vtg pada ikan koan. Data-data yang diperoleh tidak hanya melengkapi data-data biologi ikan koan dan evolusinya berdasarkan protein Vtg, namun juga dapat menunjang pengembangan metode manipulasi reproduksi pada ikan koan di masa depan.

METODE PENELITIAN

Analisis Filogenetik

Analisis filogenetik dilakukan menggunakan sekuen asam amino Vtg dari sejumlah spesies yang diperoleh dari GenBank (https://www.ncbi.nlm.nih.gov). Ketiga sekuen asam amino Ci-Vtg yang digunakan antara lain Vtg1 (APB93352.1), Vtg2 (APB93353.1), dan Vtg3 (APB93355.1). Sebagai pembanding, digunakan sekuen asam amino dari gen Vtg pada berbagai spesies ikan, antara lain Fundulus heteroclitus, Melanogrammus aeglefinus, Danio rerio, Cyprinus carpio, Morone americana, Hippoglossus hippoglossus, Larimichthys crocea, Acanthogobius favimanus, Sillago japonica, Verasper moseri, Gambusia affinis, dan Pagrus major. Sekuen-sekuen asam amino tersebut kemudian dilakukan penjajaran sekuen

p-ISSN: 1693-2617 e-ISSN: 2528-7613

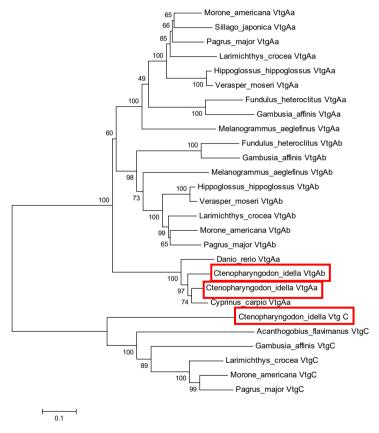
menggunakan software MEGA6 untuk pembuatan pohon filogenetik dengan metode Neighbor Joining dengan pengulangan bootstrap 1000 dengan tujuan untuk mengetahui hubungan evolusioner berdasarkan protein Vtg.

Analisis Protein

Analisis karakter fisikokimia dari protein Ci-VtgAa (APB93352.1), (APB93353.1), dan Ci-VtqC (APB93355.1) dilakukan dengan menggunakan web server ExPASy (https://web.expasy.org/protparam/). Analisis situs fosforilasi dilakukan dengan menggunakan web server KinasePhos (kinasephos.mbc.nctu.edu.tw). Analisis conserved dilakukan menggunakan NCBI Conserved Domain www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Prediksi situs N-alikosilasi dilakukan menggunakan web server YinOYang (http://www.cbs.dtu.dk/services/ YinOYang/).

HASIL DAN PEMBAHASAN

Studi in silico telah dilakukan pada ketiga protein dari Ctenopharyngodon idella yang sekuennya diambil dari GenBank (Ci-VtgAa, APB93352.1; Ci-VtgAb, APB93353.1; dan Ci-VtgC, APB93355.1). Studi ini dilakukan untuk mengkarakterisasi protein Vtg berdasarkan properti dan komposisi protein, situs fosforilasi, conserved domain, situs N-Glycosylation dan O-Glycosylation, serta prediksi struktur protein sekunder dan tersier. Selain itu, juga dilakukan analisis filogenetik untuk mengetahui hubungan evolusioner pada tiga protein Vtg dari ikan koan dan gen lainnya (Gambar 1). Hasil analisis filogenetik menunjukkan bahwa Ci-VtgAa masih berada dalam cluster yang sama dengan Ci-VtgAb.

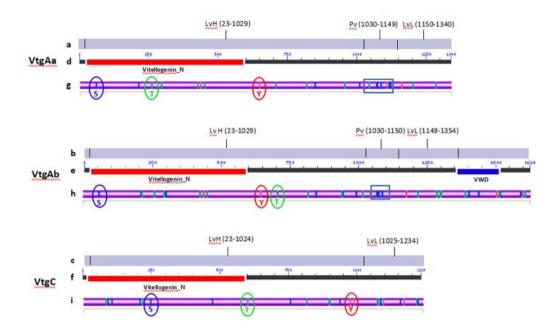


Gambar 1. Pohon filogenetik dengan metode Neighbor-joining (NJ) dari tiga protein Ci-Vtg yang dibuat dari sekuen asam amino. Pohon filogenetik dibuat dengan menggunakan software MEGA6 (Informer Technologies, Inc. New York, USA) dengan parameter Multiple Sequence Alignment antara lain: gap opening penalty=10,0; gap extension penalty=0,2; delay divergent cutoff=30%; gap separation distance=4; dan protein weight

p-ISSN: 1693-2617 e-ISSN: 2528-7613

matrix=BLOSUM. Nilai bootstrap dihitung dari 1.000 replikasi. Ketiga protein Ci-Vtg ditunjukkan dengan kotak berwarna merah.

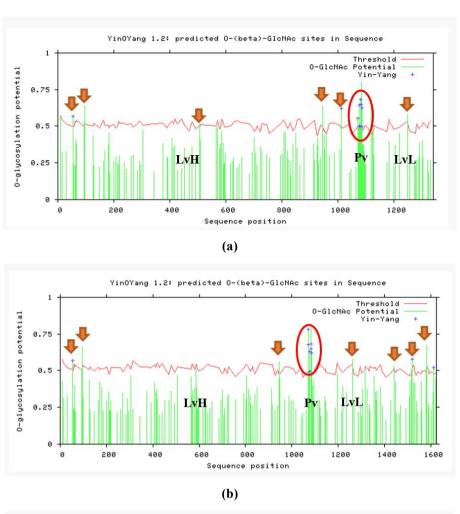
Hasil analisis domain protein yolk menggunakan Multiple Sequences Alignment menunjukkan bahwa Ci-VtgAa dan Ci-VtgAb masing-masing memiliki domain LvH, Pv, dan LvL, namun pada Ci-VtgC tidak terdapat domain Pv (Gambar 2 a, b, c). Dengan menggunakan CD-Search untuk menemukan domain conserved dari protein, telah diidentifikasi dua domain conserved pada ketiga Ci-Vtg, yaitu Vitellogenin-N, sementara VWD hanya terdapat pada Ci-VtgAb (Gambar 2 e, f, g). Lebih lanjut, situs fosforilasi juga diidentifikasi menggunakan tool online. Ditemukan daerah yang kaya akan Serine pada domain Pv (Gambar 2 g, h, i).

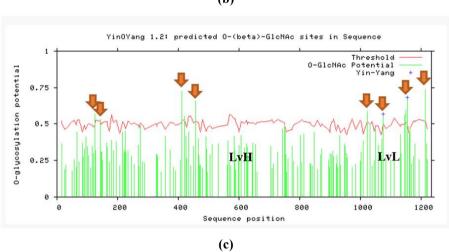


Gambar 2. Prediksi domain protein dan situs modifikasi pasca-transkripsi. (a), (b), dan (c) menunjukkan domain protein yolk dari ketiga Ci-Vtg. Ci-VtgAa dan Ci-VtgAb mengandung domain Pv, sedangkan pada Ci-VtgC tidak terdapat domain Pv. (d), (e), dan (f) menunjukkan sekuen conserved, dimana pada Ci-VtgAb terlihat adanya homologi dengan faktor von Willebrand (VWD). Strip merah merepresentasikan domain superfamili Vitellogenin-N (Accesion number: pfam1347), sedangkan strip biru menrepresentasikan superfamili VWD (Accesion number: smart00216). (g), (h), dan (I) menunjukkan fosforilasi threonine (T), tyrosine (Y), dan serine (S). Kotak biru menunjukkan fosforilasi yang terjadi pada domain Pv.

Analisis O-glikosilasi pada ketiga protein dilakukan dengan menggunakan web server YinOYang dengan hasil yang ditunjukkan pada Gambar 3. Pada protein Ci-VtgAa dan Ci-VtgAb, terjadi O-glikosilasi pada domain Pv, sedangkan pada protein Ci-VtgC O-glikosilasi terjadi pada domain LvH dan LvL.

p-ISSN: 1693-2617 e-ISSN: 2528-7613





Gambar 3. Prediksi situs O glikosilasi pada protein (a) Ci-VtgAa; (b) Ci-VtgAb; dan (c) Ci-VtgC. Panah merah menunjukkan situs glikosilasi, sedangkan oval merah menunjukkan glikosilasi yang terjadi pada domain Pv.

Untuk mengetahui karakter fisikokimia dari ketiga protein Ci-Vtg, dilakukan analisis dengan menggunakan tool ProtParam pada web server ExPASy (Tabel 1). Nilai isoelectric point (pI) dari Ci-VtgAa sebesar 8,64 dan Ci-VtgAb sebesar 7,84 yang mengindikasikan bahwa molekul-molekul tersebut bersifat basa, sementara Ci-VtgC sebesar 6,47 yang mengindikasikan bahwa protein bermuatan negatif dan dapat mengendap dalam medium

p-ISSN: 1693-2617 e-ISSN: 2528-7613

yang asam. Parameter lain yaitu instability index (II) yang mengindikasikan stabilitas protein ketika berada di laboratorium. Bila nilai II kurang dari 40, maka bisa dikatakan bahwa protein tersebut bersifat stabil. Nilai II dari Ci-VtgAa sebesar 44,59, Ci-VtgAb sebesar 42,06 dan Ci-VtgC sebesar 38,67. Data tersebut menggambarkan bahwa protein Ci-VtgC lebih stabil dibandingkan protein Ci-VtgAa dan Ci-VtgAb. Aliphatic index dari protein Ci-VtgAa sebesar 108,57 yang mengindikasikan bahwa protein ini tahan terhadap rentangan temperatur yang luas, begitu juga dengan nilai aliphatic index protein Ci-VtgAb sebesar 104,51 dan Ci-VtgC sebesar 93,28. Nilai Grand Average of Hydropathicity (GRAVY) dari Ci-VtgAa sebesar 0,148 yang mengindikasikan interaksi yang baik dengan molekul air, begitu juga dengan pada nilai GRAVY Ci-VtgAb sebesar 0,078 dan Ci-VtgC sebesar -0,129. Pada Ci-VtgAa, asam amino dengan persentase terbesar adalah Alanin (14,3%) dan Leucine (11,1%), begitu juga dengan Ci-VtgAb yaitu Alanin (12,3%) dan Leucine (10,4%). Sementara itu, pada Ci-VtgC persentase asam amino terbesar adalah Leucine (10,5%).

Tabel 1. Karakter Fisikokimia

	Ci-VtgAa	Ci-VtgAb	Ci-VtgC
Panjang asam amino	1.340	1.624	1.234
Berat molekuler (kDa)	145.660,07	178.444,02	137.871,79
Isoelectric point (pI)	8,64	7,84	6,47
Instability index (II)	44,59	42,06	38,67
Aliphatic index	108,57	104,51	93,28
Grand Average of	0,148	0,078	-0,129
Hydropathicity (GRAVY)			
Asam amino terbanyak	Alanin dan Leucine	Alanin dan Leucine	Leucine

Hasil analisis filogenetik terhadap ketiga protein Ci-Vtg menunjukkan bahwa Ci-VtgAa masih berada dalam cluster yang sama dengan Ci-VtgAb. Hal tersebut sesuai dengan studi sebelumnya yang menyebutkan bahwa sekuen protein VtgAa dan VtgAb pada Larimichthys crocea terdiri dari tiga domain utama yaitu Lipovitellin Heavy (LvH), Phosvitin (Pv), Lipovitellin Light (LvL). Berbeda dengan kedua protein tersebut, VtgC cenderung tidak mengandung domain Pv (Gao et al., 2019). Oleh karena itu, Ci-VtgC terletak pada cluster yang berbeda dan terletak pada cluster yang sama dengan VtgC dari spesies lain (Gambar 1). Adanya homologi antara Ci-VtgAa dan Ci-VtgAb yang lebih tinggi dibandingkan dengan Ci-VtgC mengindikasikan bahwa Ci-VtgAa dan Ci-VtgAb memiliki hubungan evolusioner yang lebih dekat dan keduanya merupakan derivat dari VtgA (Finn & Kristoffersen, 2007). Vtg dihasilkan oleh hati kemudian ditransfer melalui aliran darah ke ovum saat maturasi ovum. Vtg kemudian dipecah menjadi berbagai protein yolk oleh cathepsin D, dan protein yolk ini akan mengalami degradasi sehingga membentuk Fatty Acid Activation (FAA) untuk menginduksi hidrasi oosit sehingga oosit dapat mengapung dalam air (Reading & Sullivan, 2011).

Pada hewan vertebrata, termasuk ikan, Vtg mengandung empat domain yang akan terpecah menjadi sejumlah protein yolk, antara lain Lipovotelin 1/LvH, Phosvitin, Lipovitelin 2/LvL, serta domain von Willebrand tipe D (YGP40) (Finn, 2007). Domain LvH berperan dalam transpor lipid dengan membentuk "keranjang" dengan lumen residu hidrofobik (Reading & Sullivan, 2011). Domain Pv yang bersifat hidrofilik, berkontribusi dalam mempertahankan solubilitas dari protein Vtg. Adanya sejumlah Serine dalam domain Pv mengindikasikan terjadinya Fosforilasi sehingga memicu stabilitas dan hidrofilik dari Vtg, dan fosforilasi Serine mendorong terjadinya pengikatan Ca2+, Fe3+, Mg2+, Zn2+ dan kation logam lainnya (Finn, 2007; Reading & Sullivan, 2011). Sejumlah kation logam ini penting

p-ISSN: 1693-2617 e-ISSN: 2528-7613

untuk ketahanan hidup dari embrio, terutama pada ikan air tawar yang tidak dapat mengambil ion logam dari lingkungannya (Reading & Sullivan, 2011).

Hasil analisis O-glikosilasi pada protein Ci-VtgAa dan Ci-VtgAb menunjukkan bahwa terjadi O-glikosilasi pada domain Pv, sedangkan pada protein Ci-VtgC O-glikosilasi terjadi pada domain LvH dan LvL. O-glikosilasi merupakan pengikatan Glycans pada Serine dan Threonine. Terjadinya O-glikosilasi mendorong terjadinya penyediaan karbohidrat untuk oosit yang sedang berkembang (Reading & Sullivan, 2011).

PENUTUP

Hasil karakterisasi protein Vtg pada ikan koan (Ctenopharyngodon idella) menunjukkan bahwa Ci-VtgAa dan Ci-VtgAb memiliki hubungan evolusioner yang lebih dekat dibanding dengan Ci-VtgC. Hal tersebut didukung dengan analisis prediksi domain protein yolk, dimana Ci-VtgC tidak memiliki domain Phosvitin (Pv). Hasil analisis karakter fisikokimia mengindikasikan bahwa protein Ci-VtgC cenderung bersifat asam namun lebih stabil dibandingkan Ci-VtgAa dan Ci-VtgAb. Ketiga protein memiliki ketahanan terhadap rentangan temperatur yang luas, memiliki interaksi yang baik dengan molekul air, serta komposisi asam aminonya didominasi oleh Leucine (Leu). Adanya situs fosforilasi dan O-glikosilasi pada ketiga protein mengindikasikan keterlibatan peran Ci-VtgAa, Ci-VtgAb, dan Ci-VtgC dalam transpor ion logam dan karbohidrat. Hasil analisis karakterisasi secara in-silico yang dilakukan terhadap ketiga protein Vtg diharapkan dapat menambah teori dasar terkait reproduksi ikan koan.

UCAPAN TERIMA KASIH

Para penulis mengucapkan terima kasih kepada X.Y. Cheng dan X.F. Liang yang telah mempublikasikan sekuen utuh dari protein VtgAa, VtgAb, dan VtgC dari ikan koan sehingga ketiga sekuen tersebut dapat dianalisis dalam penelitian ini.

DAFTAR KEPUSTAKAAN

- Babo, D., Sampekalo, J., & Pangkey, H. (2013). Pengaruh beberapa jenis pakan hijauan terhadap pertumbuhan ikan Koan Stenopharyngodon idella. Budidaya Perairan, 3, 1-6. doi: 10.35800/bdp.1.3.2013.2716
- Choirom, R., Syahrizal., & Safratilofa. (2020). Pemberian Pakan Nabati Tumbuhan Air Yang Berbeda Terhadap Pertumbuhan dan Kelangsungan Hidup Benih Ikan Koan (Ctenopharyngodon idella). Jurnal Akuakultur Sungai dan Danau, 5(1), 24-29. doi: 10.33087/akuakultur.v4i2.55
- Cudmore, B. & Mandrak, N.E. (2004). Biological Synopsis of Grass Carp (Stenopharyngodon idella): Canadian Manuscript Report of Fisheries and Aquatic Sciences 2705. Burlington-Kanada.
- Finn, R.N. (2007). Vertebrate yolk complexes and the functional implications of phosvitins and other subdomains in vitellogenins. Biol Reprod 76(6),926–935. doi: 10.1095/biolreprod.106.059766
- Finn, R.N. & Kristoffersen, B.A. (2007). Vertebrate vitellogenin gene duplication in relation to the "3R hypothesis": correlation to the pelagic egg and the oceanic radiation of teleosts. PLoS One, 2(1), 169. doi: 10.1371/journal.pone.0000169
- Gao, X., Zhou, Y., Zhang, D., Hou, C. & Zhu, J. (2019). Multiple vitellogenin genes *vtgs) in large yellow croaker (Larimichthys crocea): molecular characterization and expression pattern analysis during ovarian development. Fish Physiol Biochem. https://doi.org/10.1007/s10695-018-0569-y
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D. & Bairoch, A.

p-ISSN: 1693-2617 e-ISSN: 2528-7613

- (2005). Protein Identification and Analysis Tools on the ExPASy Server. (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press pp. 571-607. doi: 10.1385/1-59259-584-7:531
- Kementerian Kelautan dan Perikanan. (2018). Produktivitas Perikanan Indonesia. Forum Merdeka Barat 9 Kementerian Komunikasi dan Informatika.
- Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C.J., Lu, S., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Lu, F., Marchler, G.H., Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang, D., Zheng, C., Geer, L.Y. & Bryant, S.H. (2017). CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. Nucleic Acids Res, 45, 200-203. doi: 10.1093/nar/qkw1129
- Pipil, S., Rawat, V.S., Sharma, L. & Sehgal, N. (2015). Characterization of incomplete vitellogenin (VgC) in the indian freshwater murrel, Channa punctatus (Bloch). Fish Physiol Bioche, 41(1), 107. doi: 10.1007/s10695-014-0009-6
- Reading, B. & Sullivan, C. (2011). Vitellogenesis in fishes. In: Ferrell AP (ed) Encyclopedia of fish physiology: from genome to environment. Academic Press, Inc, Maryland Heights: MO, pp 635–646.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution, 30, 2725-2729. doi: 10.1093/molbev/mst197
- Wong, Y.H., Lee, T.Y., Liang, H.K., Huang, C.M., Wang, T.Y., Yang, Y.H., Chu, C.H, Huang, H.D., Ko, M.T. & Hwang, J.K. (2007). KinasePhos 2.0: A web server for identifying protein kinase-specific phosphorylation sites based on sequences and coupling patterns. Nucleic Acids Research. doi: 10.1093/nar/gkm322