Studying the Features of Proteins Surrounding PLC Gama During Fertilization by Using Bioinformatics Databases

Rana Hussein Naser

Department of Biotechnology – College of Science – University of Diyala

ranaalqaysi@gmail.com

Received: 21 October 2018 **Accepted: 24 January 2019**

Abstract

Fertilization simply is about upstream of kinases and protein interactions, so the justified bioinformatics tools were used to understand and analyze the interactions between proteins during fertilization progress, that's would help a lot in this research to understand the protein-protein interaction. The whole mark of the experiment at hand is to reveal bioinformatics data related to PLC gamma and utilize the collected data to understand the concept of fertilization. The bioinformatics tools were performed to generate a useful biological knowledge.

Keywords: Fertilization, Bioinformatics Databases, Plc Gamma, protein interaction.

دراسة خصائص البروتينات المحيطة بـجزيئة PLC GAMA خلال عملية الاخصاب باستخدام قاعدة بيانات المعلوماتية الحيوية

رنا حسين ناصر

قسم التقنيات الاحيائية – كلية العلوم – جامعة ديالى

Studying the Features of Proteins Surrounding PLC GAMA DURING FERTILIZATION by Using Bioinformatics Databases

Rana Hussein Naser

ألخلاصة

الإخصاب ببساطة هو منبع من اثر الكاينيزات و التحركات والتفاعلات البروتينية، لذلك أردنا استخدام ألادوات المعلوماتية الحيوية لفهم وتحليل التفاعلات بين البروتينات خلال الإخصاب، وهذا من شأنه أن يساعد كثيرا في الكشف عن البيانات المعلوماتية الحيوية المتعلقة PLC كاما والاستفادة من البيانات التي تم جمعها لفهم مفهوم الإخصاب. لذا نأمل هنا استخدام أدوات المعلوماتية الحيوية لتوليد المعرفة البيولوجية المفيدة.

الكلمات المفتاحية: الاخصاب، بيانات المعلومات الحيوية، الجزيئية كاما PLC، التفاعلات البروتينية.

Introduction

Fertilization in humans involves the fusion of an egg and a sperm in the ampulla of the uterine tube. This usually results into the formation of the zygote. When the acrosome of the sperm encounters an ovum, it secretes enzymes which assist it to burrow into the jelly coat covering the egg. There is a fusion between sperm's and egg's plasma membranes; the fertilized egg then moves to the uterus where the zygote develops. Research has shown that egg activation (transition of an egg to an embryo) requires an increase in Ca^{2+} [4]. There are several mechanisms that are associated with the elevation of $Ca2^{+}$, which are required for egg activation. For instance, in some animals, the production of Inositoal 1, 4, 5 triphosphate (IP3), the participation of phospholipase C and the subsequent activation of IP3 receptors [4,10].

Generally, egg activation contributes to the re-entry of the egg into the cell cycle and the subsequent embryonic development. It should be noted that egg activation is highly influenced by the increase in Ca^{2+} concentration in the cytoplasm [4]. As such, having a detailed understanding of how an increase in the concentration of cytoplasmic Ca^{2+} is the first step towards understanding the concept of fertilization.

In ascidians and echinoderms, there are distinct signaling pathways which determine the sequence of events that occur beginning from the interaction of the egg and sperm to the release of Ca^{2+} from the egg endoplasmic reticulum [12]. One of the hypothesized models involves the activation of Src family kinases (SFK) of the egg which often leads to the phosphorylation of phospholipase C. During this sequence, phosphoinositide biphosphate (PIP2) is hydrolyzed to

Inositoal 1, 4, 5 triphosphate (IP3) which stimulates the production of intracellular Ca^{2+} [12]. This is attained via the IP3 receptor Ca^{2+} in the endoplasmic reticulum of the egg. The hydrolysis of PIP2 to IP3 is catalyzed by Phospholipase C (PLP) family members. In echinoderms, PLC-beta and PLC gamma are the most common PLC family members that are utilized [12]. PLC gamma has been found to be activated by receptor and non-receptor tyrosine protein kinases (Tpk). In addition, PLC gamma can be activated through its translocation to the plasma membrane where PIP2 is chiefly found [12].

Materials and Methods

1. PLC γ cDNA sequence

The current study we utilized starfish *Asterina miniata* sp. PLCγ cDNA. The homology of the protein of this species (AmPLCγ) with mammalian PLCγ1 is 49%. The AmPLCγ exhibited a recombination with 58-kDa Src family kinase and this interrelation forms recombinant AmPLCγ Src homology 2(SH2) domains which particularly helps in fertilization. gi|40365362|gb| accession number /AY486068.1/ Asterina miniate phospholipase C-γ MRNA, complete cds.

2. Identification and characterization of proteins

The String database (http://string-db.org/) was used to identify known as well as predicted proteins from all the organisms which are showing interactions with PLCγ1 starfish sequence. Based on maximum homology with starfish PLC γ1 sequence. By using STRING database, we have screened the protein of interest according to maximum homology score. In general *Homo sapiens* sp. was identified as a most suitable organism for study because their protein exhibited maximum homology scores. The percentage of cDNA homology score vary from 26-58% among various organisms including *Homo sapiens.*

3. BLAST and its use in study

To find the regions of local similarity between PLC γ1 starfish and other sequences. BLAST tool was used. By using tBLASTn translated nucleotide sequence i.e., *Asterina miniata* phospholipase C-γ mRNA, complete cds compare with protein sequences of interest from databases and calculates the statistical significance of matches. Thus, by using tBLASTn for this study we can infer functional and evolutionary relationships between starfish and other organism sequences as well as help to identify members of gene families.

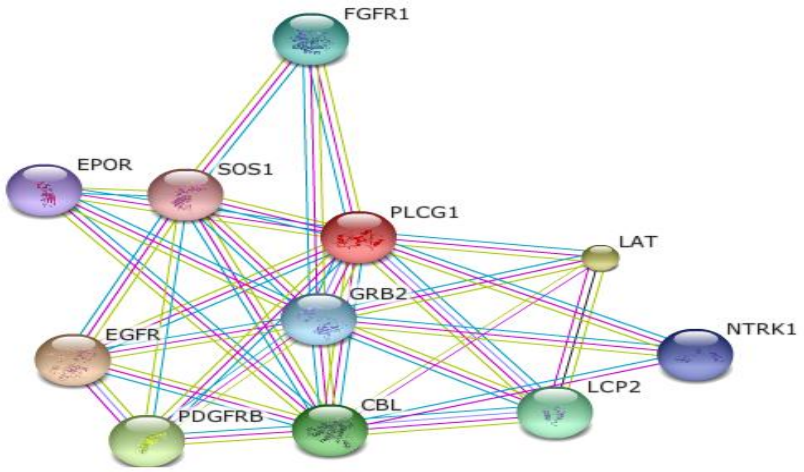
Results and Discussion

STRING database is a biological database of known and predicted protein-protein interaction. It is freely accessible and regularly updated. It is a very important tool to analyze protein's interactions with other surrounding proteins, which is the most significant aim for this project. Also, it allows analyzing the protein in all other orthologous. STRING database is an interesting tool where gives protein-protein interaction network to visualize by. This network is available with many views that can help more to know if the surrounding protein is strongly associated with the interest protein or not. Furthermore, it allows studying each surrounded protein in detail and collect all its available information about to build up an own database like a definition of the protein and its sequence and the domains of this protein and some information about each domain.

The goal of these experiments was to identify proteins that may represent targets for interaction with PLCγ during fertilization in the starfish. The STRING database (http://string-db.org/) was used to identify known as well as predicted proteins that interact with PLCγ1 in other systems. Attention was focused on human PLCγ1 because Homo sapiens has been heavily studied and it is an ideal system to identify proteins which bind with PLCγ1 that have homology to a starfish sequence. The sequences of starfish which are showing homology will be compared to the corresponding human sequence to get a complete comparative analysis of proteins of this system. Based on the homology modeling the results of identified proteins that show binding to PLCγ and homology with starfish system

Studying the Features of Proteins Surrounding PLC GAMA DURING FERTILIZATION by Using Bioinformatics Databases

Rana Hussein Naser



(http://string-db.org/)

Figure 1: Proteins showing interaction with PLCγ1 in *Homo sapiens*, different line color represents the different type of association and this picture is adapted from string database (http://stringdb.org/newstring_cgi/show_network_section.pl).

In this study we will discuss two proteins, NTRK1 is showing low homology with starfish sequence while FGFR1 is showing high homology score with starfish sequence.

NTKR1 (Neurotrophic tyrosine kinase receptor) Protein

Neurotrophic tyrosine kinase receptor (NTKR) family proteins are interacting and binding with membranes and this interaction leads to auto-phosphorylation as well as for downstream proteins like those of the MAP Kinase pathway. This protein has various important roles including cell differentiation and participating in specification of sensory neuron subtypes.

Function of protein

The NTKR1 protein is critically involved in growth and development of nervous system (central as well as peripheral). The primary role is in growth, propagation, and persistence of sympathetic nervous neurons [11,5] reported that Congenital Insensitivity to Pain with Anhidrosis (CIPA) occurred due to mutations in NTRK1 gene

Studying the Features of Proteins Surrounding PLC GAMA DURING FERTILIZATION by Using Bioinformatics Databases

Rana Hussein Naser

PLCγ1 interact with NTKR1

NGF, as well as the NTRK1 protein, play a major role in neuronal cell growth and differentiation. The NGF-NTRK1 complex has the ability to phosphorylate proteins on specific tyrosine residues and thus achieve the ability to interact with proteins of signal-transduction pathways for example phospholipase C (PLC-γ), phosphatidylinositol-3'-kinase (PI3'-K) and SHC [7]. In *Homo sapiens* using string database it was identified that NTRK1 and PLCγ1 displayed effective protein-protein interaction with binding score of 0.983.

MLRGGRRGQLGWHSWAAGPGSLLAWLILASAGAAPCPDACCPHGSSGLRCTRDGALDSLHHLPGAENLTELYIENQQHLQ
HLELRDLRGLGELRNLTIVKSGLRFVAPDAFHFTPRLSRLNLSFNALESLSWKTVQGLSLQELVLSGNPLHCSCALRWLQ
RWEEEGLGGVPEQKLQCHGQGPLAHMPNASCGVPTLKVQVPNASVDVGDDVLLRCQVEGRGLEQAGWILTELEQSATVMK
SGGLPSLGLTLANVTSDLNRKNVTCWAENDVGRAEVSVQVNVSFPASVQLHTAVEMHHWCIPFSVDGQPAPSLRWLFNGS
VLNETSFIFTEFLEPAANETVRHGCLRLNQPTHVNNGNYTLLAANPFGQASASIMAAFMDNPFEFNPEDPIPVSFSPVDT
NSTSGDPVEKKDETPFGVSVAVGLAVFACLFLSTLLLVLNKCGRRNKFGINRPAVLAPEDGLAMSLHFMTLGGSSLSPTE
GKGSGLQGHIIENPQYFSDACVHHIKRRDIVLKWELGEGAFGKVFLAECHNLLPEQDKMLVAVKALKEASESARQDFQRE
AELLTMLQHQHIVRFFGVCTEGRPLLMVFEYMRHGDLNRFLRSHGPDAKLLAGGEDVAPGPLGLGQLLAVASQVAAGMVY
LAGLHFVHRDLATRNCLVGQGLVVKIGDFGMSRDIYSTDYYRVGGRTMLPIRWMPPESILYRKFTTESDVWSFGVVLWEI
FTYGKQPWYQLSNTEAIDCITQGRELERPRACPPEVYAIMRGCWQREPQQRHSIKDVHARLQALAQAPPVYLDVLG

Figure 2: NTRK1 protein sequence

Studying the Features of Proteins Surrounding PLC GAMA DURING FERTILIZATION by Using Bioinformatics Databases

Rana Hussein Naser

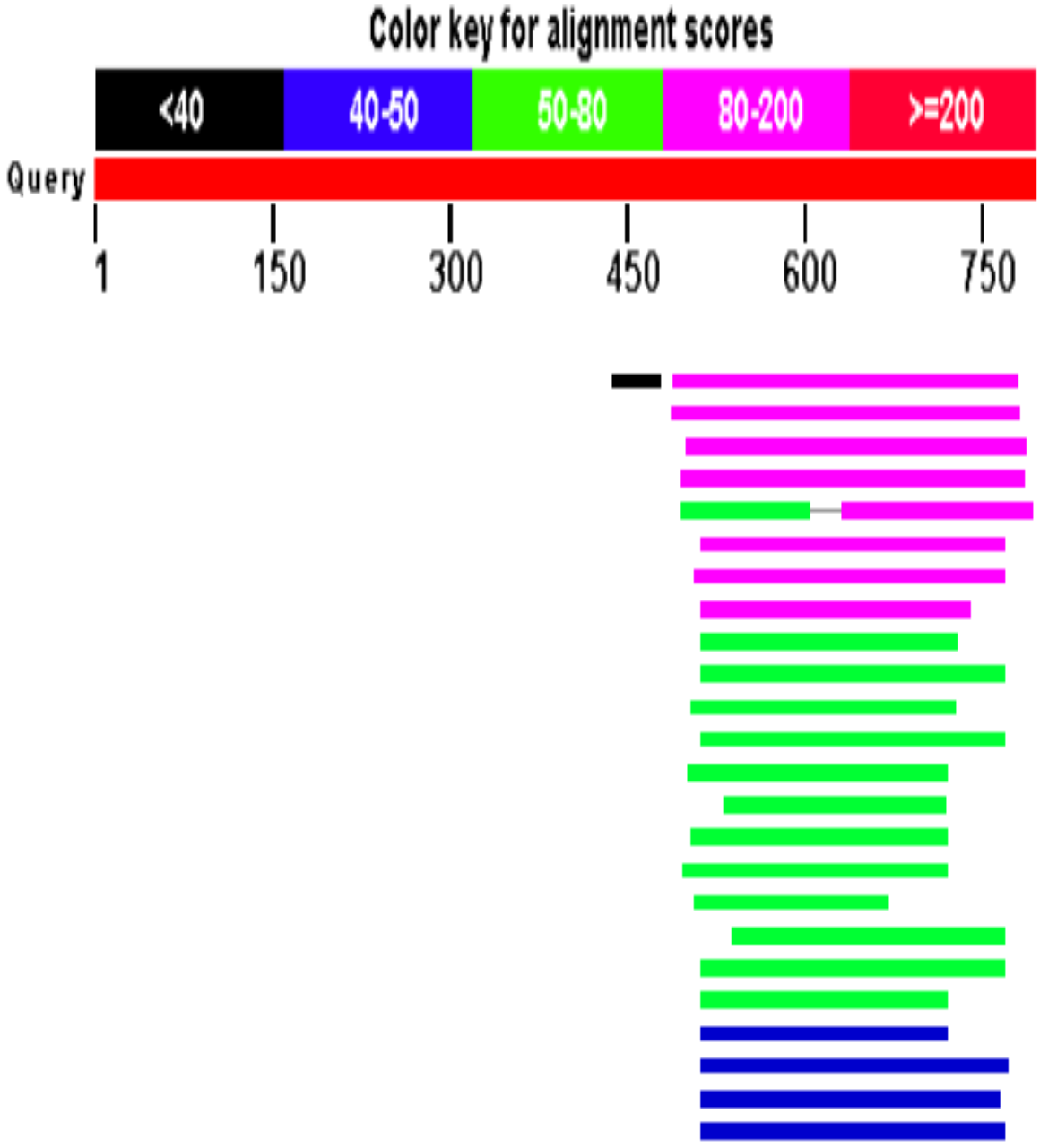


Figure 3: Graphic overview showing score distribution of 26 BLAST hits on the NTRK1 protein sequence with highest (80-200) alignment score for *Asterina miniata* Src family tyrosine kinase (SFK1) mRNA, complete cds (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Studying the Features of Proteins Surrounding PLC GAMA DURING FERTILIZATION by Using Bioinformatics Databases

Rana Hussein Naser

Description	Max score	Total score	Query cover	E value	Ident	Accession
Asterina miniata Src family tyrosine kinase (SFK1) mRNA, complete cds	179	179	36%	1e-48	35%	AY518774.1
Asterina miniata Src family kinase (SFK2) mRNA, complete cds	177	177	37%	6e-48	35%	AY518775.1
Asterina miniata Src family kinase (SFK3) mRNA, complete cds	174	174	36%	2e-46	36%	AY518776.1
Asterina miniata C-terminal Src kinase (CSK) mRNA, complete cds	166	166	36%	1e-43	34%	AY518773.1
Patiria pectinifera vegfr mRNA for vascular endothelial growth factor receptor, partial cds	131	187	34%	2e-32	42%	AB705446.1
Asterina pectinifera sfRSK mRNA for p90 ribosomal S6 kinase, complete cds	97.1	97.1	32%	2e-21	29%	AB073313.1
Marthasterias glacialis partial mRNA for aurora/ipl1p-related kinase (aurora gene)	86.7	86.7	33%	3e-18	27%	AJ889572.1
Patiria pectinifera aur mRNA for aurora kinase, complete cds	84.3	84.3	28%	8e-18	29%	AB530259.1
Asterina pectinifera mRNA for cdc2 (possible component of maturation-promoting factor), complete cds	73.2	73.2	27%	2e-14	27%	D79982.1
Pisaster ochraceus protein kinase C-related kinase (PRK2) mRNA, partial cds	68.2	68.2	32%	3e-12	28%	AF035554.1
Patiria pectinifera protein kinase C isoform (nPKC) mRNA, complete cds	62.4	62.4	28%	1e-10	28%	FJ826887.1
Patiria pectinifera protein kinase C isoform (cPKC) mRNA, complete cds	62.4	62.4	32%	1e-10	24%	FJ826889.1
Patiria pectinifera Cdk2 mRNA for cyclin-dependent kinase 2, complete cds	61.6	61.6	27%	2e-10	26%	AB481376.1
Marthasterias glacialis (cdk7) mRNA, partial cds	57.0	57.0	23%	1e-09	26%	U29685.1
Patiria pectinifera PDK1 mRNA for phosphoinositide dependent kinase-1, complete cds	56.6	56.6	27%	7e-09	21%	AB110536.1
Asterina pectinifera mRNA for Mos, complete cds	55.8	55.8	28%	1e-08	25%	AB040102.1
Patiria pectinifera gwl mRNA for greatwall kinase, complete cds	55.5	55.5	20%	2e-08	25%	AB597032.1
Asterias amurensis aaGC mRNA for guanylate cyclase, complete cds	53.5	53.5	29%	8e-08	21%	AB070354.1
Asterina pectinifera mRNA for kinase Akt/PKB, complete cds	53.1	53.1	32%	9e-08	24%	AB080291.1
Pisaster ochraceus maturation inhibited protein kinase p40 (MIPK) mRNA, complete cds	52.4	52.4	26%	1e-07	26%	AF084674.1
Marthasterias glacialis partial mRNA for extracellular signal-regulated protein kinase (erk gene)	49.3	49.3	26%	1e-06	25%	AJ536587.1
Patiria pectinifera Plk mRNA for polo-like kinase, complete cds	49.3	49.3	32%	1e-06	21%	AB084485.1
Asterina pectinifera Myt1 mRNA, complete cds	43.1	43.1	31%	1e-04	23%	AB080280.1
Patiria pectinifera protein kinase C isoform (aPKC) mRNA, complete cds	43.1	43.1	32%	1e-04	20%	FJ826888.1
Odontaster validus amino acid transporter (ATa) mRNA, partial cds	27.3	27.3	5%	8.5	39%	KC155343.1

Figure 4: Graphic table showing sequence producing significant alignment with NTRK1 protein along with similarity or identity score of proteins, which Showing the homology with starfish and NTRK1 protein the maximum identity score obtained for *Asteria minimata* Src family kinase (SFK1) is 35 %.

Asterina miniata Src family tyrosine kinase (SFK1) mRNA, complete cds
Sequence ID: gb|AY518774.1| Length: 1836 Number of Matches: 1

Range 1: 849 to 1661 GenBank Graphics Next Match Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
179 bits(455)	1e-48	Compositional matrix adjust.	104/296(35%)	168/296(56%)	30/296(10%)	+3

```
Query  492   ENPQYFS---DACVHHIKRRDIVLKWELGEGAFGKVFLAECHNLLPEQDKMLVAVKALKE  548
             ENP    S    DA     I R   + L+ +LG G FG+V+     +    P       VA+K LK+
Sbjct  849   ENPNTVSLGRDAW--EIPRTSLTLESKLGAGQFGEVWKGTWNGKTP------VAIKTLKK  1004

Query  549   ASESARQDFQREAELLTMLQHQHIVRFFGVCTEGRPLLMVFEYMRHGDLNRFLRS-HGPD  607
              + +     F   EA ++   L+H   + + + VC++   P+ +V E M +G L   FL+    G +
Sbjct  1005  GTMTPTA-FLAEANIMKKLRHPKLCQLYAVCSDKEPIYIVAELMCNGSLLDFLKDGEGRN  1181

Query  608   AKLLAGGEDvapgplglgqllavaSQVAAGMVYLAGLHFVHRDLATRNCLVGQGLVVKIG  667
              KL                    +L+ + +Q+A+GM +L   +++VHRDLA RN LVG+G +VK+
Sbjct  1182  LKL---------------PELVDMGAQIASGMAFLESMNYVHRDLAARNVLVGEGNIVKVA  1319

Query  668   DFGMSRDIYSTDYY-RVGGRTMLPIRWMPPESILYRKFTTESDVWSFGVVLWEIFTYGKQ  726
             DFG++R I   T+Y   R G +    PI+W   PE+ +Y +FT +SDVWSFGV+L E+ T+G+
Sbjct  1320  DFGLARMIEDTEYTARQGAK--FPIKWTAPEAAMYGRFTIKSDVWSFGVLLTELVTHGRI  1493

Query  727   PWYQLSNTEAIDCITQGRELERPRACPPEVYAIMRGCWQREPQQRHSIKDVHARLQ  782
             P+   + N E +D +   G   + +    CP   +Y +M+ CW ++P   RH+ + + +H+ L
Sbjct  1494  PYPGMMNMEVLDQVEHGYRMPKMANCPDTLYELMQKCWDKDPAARHTFEFLHSYLD  1661
```

Figure 5: Alignment statistics of *Asterina miniata* SFK1 mRNA.

FGFR1 (Fibroblast growth factor receptor) protein

Fibroblast growth factor receptor (FGFR) is a protein with highly conserved amino acid sequences between members and throughout evolution. Based on tissue distribution and ligand affinities the variation persists among members of FGFR.

Function protein

This protein plays an important role in the regulation of embryonic development, differentiation, cell proliferation and migration. It also plays a critical role in normal mesodermal patterning and correct axial organization during embryonic development, normal skeletogenesis and normal development of the gonadotropin-releasing hormone (GnRH) neuronal system

PLCγ1 interact with FGFR1

This protein is very important for activation of various signaling molecules, such as PLCγ1, whose activation leads to synthesis of diacylglycerol and inositol 1,4,5-trisphosphate. Similarly, phosphorylation of FRS2 activates recruitment of GRB2, GAB1, PIK_3R_1 and SOS_1, and mediates triggering of RAS, MAPK1/ERK2, MAPK3/ERK1 from the MAP kinase signaling

pathway, as well as of the AKT1 signaling pathway. It also stimulates phosphorylation of SHC1, STAT1 and PTPN11/SHP2. FGFR is regulated by a secondary SH2 domain binding [2]. FGF2 becomes stimulated due to its binding to FGFR1, which in turn hydrolyzes cytosolic form of PLCγ1 and Phosphatidylinositol 4,5 bisphophate to diacylglycerol (DAG) and Inositol triphosphate (IP3).

The SH2 domain of PLCγ1 represents marked association with the epidermal growth factor (EGF) or platelet derived growth factor (PDGF)-receptors in starfish. Earlier reports in bacterial system also suggest that SH2 domain displaying similar kind of interaction [1].

MWSWKCLLFWAVLVTATLCTARPSPTLPEQAQPWGAPVEVESFLVHPGDLLQLRCRLRDDVQSINWLRDGVQLAESNRTR
ITGEEVEVQDSVPADSGLYACVTSSPSGSDTTYFSVNVSDALPSSEDDDDDDDSSSEEKETDNTKPNRMPVAPYWTSPEK
MEKKLHAVPAAKTVKFKCPSSGTPNPTLRWLKNGKEFKPDHRIGGYKVRYATWSIIMDSVVPSDKGNYTCIVENEYGSIN
HTYQLDVVERSPHRPILQAGLPANKTVALGSNVEFMCKVYSDPQPHIQWLKHIEVNGSKIGPDNLPYVQILKTAGVNTTD
KEMEVLHLRNVSFEDAGEYTCLAGNSIGLSHHSAWLTVLEALEERPAVMTSPLYLEIIIYCTGAFLISCMVGSVIVYKMK
SGTKKSDFHSQMAVHKLAKSIPLRRQVTVSADSSASMNSGVLLVRPSRLSSSGTPMLAGVSEYELPEDPRWELPRDRLVL
GKPLGEGCFGQVVLAEAIGLDKDKPNRVTKVAVKMLKSDATEKDLSDLISEMEMMKMIGKHKNIINLLGACTQDGPLYVI
VEYASKGNLREYLQARRPPGLEYCYNPSHNPEEQLSSKDLVSCAYQVARGMEYLASKKCIHRDLAARNVLVTEDNVMKIA
DFGLARDIHHIDYYKKTTNGRLPVKWMAPEALFDRIYTHQSDVWSFGVLLWEIFTLGGSPYPGVPVEELFKLLKEGHRMD
KPSNCTNELYMMMRDCWHAVPSQRPTFKQLVEDLDRIVALTSNQEYLDLSMPLDQYSPSFPDTRSSTCSSGEDSVFSHEP
LPEEPCLPRHPAQLANGGLKRR

Figure 6: FGFR1 protein sequence

Studying the Features of Proteins Surrounding PLC GAMA DURING FERTILIZATION by Using Bioinformatics Databases

Rana Hussein Naser

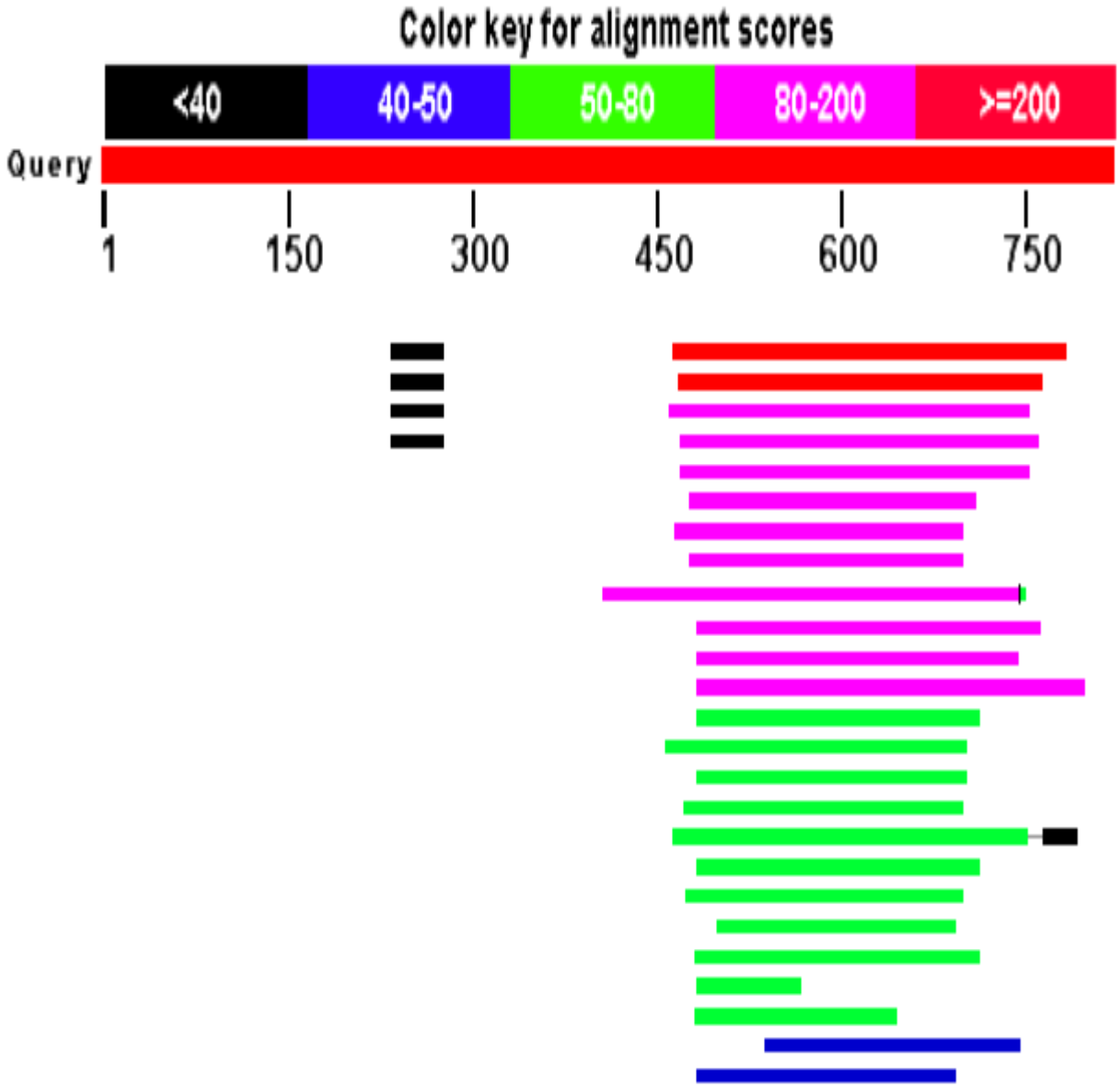


Figure 7: Graphic overview showing score distribution of 29 BLAST hits on the FGFR protein sequence with highest (80-200) alignment score for *Patiria pectinifera* vegfr mRNA (http://blast.ncbi.nlm.nih.gov/Blast.cgi)

Studying the Features of Proteins Surrounding PLC GAMA DURING FERTILIZATION by Using Bioinformatics Databases

Rana Hussein Naser

Description	Max score	Total score	Query cover	E value	Ident	Accession
Patiria pectinifera vegfr mRNA for vascular endothelial growth factor receptor, partial cds	317	317	38%	2e-97	45%	AB705446.1
Asterina miniata Src family kinase (SFK2) mRNA, complete cds	201	201	35%	5e-56	40%	AY518775.1
Asterina miniata Src family kinase (SFK3) mRNA, complete cds	196	196	35%	9e-54	38%	AY518776.1
Asterina miniata Src family tyrosine kinase (SFK1) mRNA, complete cds	183	183	35%	9e-50	37%	AY518774.1
Asterina miniata C-terminal Src kinase (CSK) mRNA, complete cds	184	184	34%	2e-49	38%	AY518773.1
Patiria pectinifera protein kinase C isoform (nPKC) mRNA, complete cds	93.2	93.2	28%	3e-20	31%	FJ826887.1
Patiria pectinifera aur mRNA for aurora kinase, complete cds	88.2	88.2	28%	5e-19	30%	AB530259.1
Marthasterias glaciallis partial mRNA for aurora/ipl1p-related kinase (aurora gene)	88.6	88.6	27%	9e-19	30%	AJ889572.1
Asterina pectinifera sfRSK mRNA for p90 ribosomal S6 kinase, complete cds	86.3	158	41%	5e-18	26%	AB073313.1
Asterina pectinifera mRNA for Mos, complete cds	84.0	84.0	34%	1e-17	25%	AB040102.1
Patiria pectinifera protein kinase C isoform (cPKC) mRNA, complete cds	83.6	83.6	31%	3e-17	27%	FJ826889.1
Plsaster ochraceus protein kinase C-related kinase (PRK2) mRNA, partial cds	82.4	82.4	38%	1e-16	26%	AF035554.1
Asterina pectinifera mRNA for cdc2 (possible component of maturation-promoting factor), complete cds	78.6	78.6	28%	3e-16	26%	D79982.1
Plsaster ochraceus maturation inhibited protein kinase p40 (MIPK) mRNA, complete cds	72.8	72.8	29%	5e-14	28%	AF084574.1
Marthasterias glaciallis partial mRNA for extracellular signal-regulated protein kinase (erk gene)	70.1	70.1	26%	3e-13	28%	AJ536587.1
Patiria pectinifera PDK1 mRNA for phosphoinositide dependent kinase-1, complete cds	70.1	70.1	27%	5e-13	26%	AB110536.1
Patiria pectinifera Plk mRNA for polo-like kinase, complete cds	68.2	89.3	38%	1e-12	23%	AB084465.1
Patiria pectinifera Cdk2 mRNA for cyclin-dependent kinase 2, complete cds	68.2	68.2	28%	2e-12	26%	AB481376.1
Patiria pectinifera protein kinase C isoform (aPKC) mRNA, complete cds	65.5	65.5	27%	1e-11	23%	FJ826888.1
Marthasterias glaciallis (cdk7) mRNA, partial cds	58.2	58.2	23%	2e-09	25%	U29665.1
Asterina pectinifera mRNA for kinase Akt/PKB, complete cds	57.0	57.0	28%	6e-09	24%	AB060291.1
Patiria miniata ephrin receptor mRNA, partial cds	50.8	50.8	10%	5e-08	38%	JX844800.1
Patiria pectinifera qwl mRNA for greatwall kinase, complete cds	53.5	53.5	20%	7e-08	25%	AB597032.1
Asterias amurensis aaGC mRNA for guanylate cyclase, complete cds	49.3	49.3	25%	2e-06	25%	AB070354.1
Asterina pectinifera Myt1 mRNA, complete cds	48.1	48.1	25%	3e-06	23%	AB060280.1
Parvulastra exigua voucher Ech034 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	26.9	26.9	5%	8.8	27%	EU870017.1
Parvulastra exigua voucher Ech035 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	26.9	26.9	5%	8.8	27%	EU870016.1
Parvulastra exigua voucher Ech036 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	26.9	26.9	5%	8.8	27%	EU870015.1
Macrophychaster accrescens isolate A02.10 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	26.8	26.8	5%	9.0	27%	GU227104.1

Figure 8: Graphic table showing sequences producing significant alignment with FGFR1 protein along with similarity or identity score of proteins, and maximum identity score obtained for *Patiria pectinifera* vegfr mRNA is 45 %.

Studying the Features of Proteins Surrounding PLC GAMA DURING FERTILIZATION by Using Bioinformatics Databases

Rana Hussein Naser

Patiria pectinifera vegfr mRNA for vascular endothelial growth factor receptor, partial cds
Sequence ID: dbj|AB705446.1| Length: 1860 Number of Matches: 1

Range 1: 77 to 1162 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
317 bits(813)	2e-97	Compositional matrix adjust.	166/365(45%)	227/365(62%)	48/365(13%)	+2

```
Query  465   LPEDPRWELPRDRLVLGKPLGEGCFGQVVLAEAIGLDKDKPNRVTKVAVKMLKSDATEKD  524
             LP DP+WE PR+RL LG  LG+G FG+VV A A G+DK +    T VAVKMLK +A++ +
Sbjct  77    LPYDPKWEFPRERLKLGSILGQGAFGRVVKAAAFGIDKTQ--TCTTVAVKMLKENASDVE  250

Query  525   LSDLISEMEMMKMIGKHKNIINLLGACTQDGPLYVIVEYASKGNLREYLQARRPPGL---  581
                L++E++M+  IG H N++NL+GACT+   L +IVE+ + GNL +YL+ RR   +
Sbjct  251   RKALMTELKMLTHIGPHLNVVNLMGACTKID-LLIIVEFCTHGNLSDYLRGRRQDYVVES  427

Query  582   EYCYNPSHNPE-----------------------------------------EQLSSKDL  600
             +  + P H+ +                                         E L+ KDL
Sbjct  428   KDTHQPLHHQQRLIAASLLSTGASAEPDSPGLPLDAEDEDEDDDVFTFVEKKEPLTLKDL  607

Query  601   VSCAYQVARGMEYLASKKCIHRDLAARNVLVTEDNVMKIADFGLARDIHHIDYYKKTTNG  660
             +  A+QVARGME+LASKKCIHRDLAARNVL+ +DN++KI DFGL+RDI+H   Y     G
Sbjct  608   LCFAFQVARGMEFLASKKCIHRDLAARNVLLADDNIVKICDFGLSRDIYHDPDYVTRGGG  787

Query  661   RLPVKWMAPEALFDRIYTHQSDVWSFGVLLWEIFTLGGSPYPGVPV-EELFKLLKEGHRM  719
             RLP+KWMAPE++FD++YT  SDVWSFGV +WE+F LGG+PYPGVPV EE +  LK G+RM
Sbjct  788   RLPIKWMAPESIFDKVYTSYSDVWSFGVFMWELFQLGGTPYPGVPVDEEFYNRLKNGYRM  967

Query  720   DKPSNCTNELYMMMRDCWHAVPSQRPTFKQLVEDLDRIVALTSNQEYLDLSMPLDQYSPS  779
               P +   E+Y +M +CW+    +RP F  LV  L   +     QEYLDL++P +  + S
Sbjct  968   CAPDHAPQEIYHIMLECWNTEAKERPDFSDLVIKLGDQLEANVIQEYLDLNIPFEIENAS  1147

Query  780   FPDTR  784
               P T+
Sbjct  1148  RPSTQ  1162
```

Figure 9: Alignment statistics of *Patiria pectinifera* vegfr mRNA

The present study proposed the identification and characterization of proteins (known as well as predicted) that may interact with PLC γ1 in all the model systems and showed maximum homology with the identified starfish sequence. For identification of these proteins, a bioinformatics approach was utilized. The two proteins were identified and studied in details and how they may interact and bind to plc was investigated. The Neurotrophic tyrosine kinase receptor (NTKR1) family protein kinase receptors are known to interact and bind with membranes and this interaction leads to phosphorylation of its own proteins as well as the MAP kinase pathway proteins. The proteins of this family have many important functions, in diseases along with cell differentiation and in specification of sensory neuron subtypes. The NGF and TRK gene of NTKR family participate in the control of early and late stages of Alzheimer's disease and thus participate in degeneration of this disease [3]. This protein shows a weak interaction and binding with many growth factors including NGF, neurotrophin-3 and neurotrophin-4/5 and almost no interaction with BDNF, but large interaction and strong binding

with tropomyosin receptor kinases viz., [6]. PLC γ1, along with SHC1 and PI 3-kinase interact strongly with Trk receptor of NTRK1 and behave as known substrates for this protein. The similarity search using BLAST tool NTKR1 presented a low homology with starfish sequence. Due to low homology and similarity scores this protein sequence should not be used for PCR studies in to design primer because it has less than 10 identical bases as shown in figure 5.

The fibroblast growth factor receptor (FGFR) has highly conserved amino acid sequences between members and throughout evolution. Based on tissue distribution and ligand affinities, the variation persists among members of FGFR1. The extracellular regions of this protein interact with fibroblast growth factors and this initiates downstream signaling leading to cellular differentiation and mitogenesis. This protein is involved in variety of functions including migration, differentiation, proliferation and development of embryonic cells [9] reported that FGF ligands and receptors function in skeletogenesis and this has been confirmed through comparative genetic profiling. The FGFRs are transmembrane catalytic receptors that have significant intracellular tyrosine kinase activity. This protein is very important for activation of various signaling molecules, such as PLCγ1 activation leading to the synthesis of diacylglycerol and inositol 1,4,5-trisphosphate. This protein shows 45% homology to a starfish sequence. During BLAST this protein sequence has 10 or more identical bases and due to similarity scores this protein sequence can be used for PCR studies in form of primer.

Conclusion

The STRING database was used in this study to find and characterize proteins that may interact with PLC γ1 in all the systems and that show maximum homology with the starfish sequence. The identified proteins have a homology to a starfish sequence and could be expected to bind with PLC γ. Based on maximum similarity score during tBLASTn and showing more than 10 similar bases during homology studies, suitable proteins that can be utilized for further studies, such as in the form of primers for PCR studies present as the best target proteins.

References

1. Anderson, D.; Koch, C. A.; Grey, L.; Ellis, C.; Moran, M. F.; Pawson T. Science 1990, *250*(4983), 979-982
2. Bae, J. H.; Lew, E. D.; Yuzawa, S.; Tome, F.; Lax, I.; Schlessinger, J. *Cell 2009*, *138*(3), 514-524.
3. Counts, S. E.; Nadeem, M.; Wuu, J.; Ginsberg, S. D.; Saragovi, H. U.; Mufson, E. J. *Official Journal of the American Neurological Association and the Child Neurology Society* 2004, 56(4), 520-531.
4. Giusti, F.; Xu, W.; Hinkles, B.; Terasaki, M.; Jaffe, L. *The Journal of Biological Chemistry* 2000, 275(22), 16788-16794.
5. Indo, Y.; Tsuruta, M.; Hayashida, Y.; Karim, M.A.; Ohta, K.; Kawano, T.; Mitsubuchi, H.; Tonoki, H.; Awaya, Y.; Matsuda, I. *Nature genetics* 1996, *13*(4),485-488.
6. Ip, N. Y.; Stitt, T. N.; Tapley, P.; Klein, R.; Glass, D. J.; Fandl, J.; Greene, L. A.; Barbacid, M.; Yancopoulos, G. D. Neuron 1993,10(2), 137-149.
7. Obermeier, A.; Bradshaw, R. A.; Seedorf, K.; Choidas, A.; Schlessinger, J.; Ullrich, A. *The EMBO Journal* 1994, 13(7) ,1585–1590.
8. O'Neill, F. J.; Gillett, J.; Foltz, K. R. *Journal of Cell Science 2004*, 117(25), 6227-6238.
9. Quarto, N.; Behr, B.; Li, S.; Longaker, M. T. Cells Tissues Organs 2009,190(3),158–169.
10. Runft, L.; Mehlmann, L. *Developmental Biology* 2002, 245(2), 237-254.
11. Snider, W. D. Cell 1994, 77(5), 627–638.
12. Townley, I. K.; Schuyler, E. K.; Parker-Gur, M.; Foltz, K. R. Developmental Biology 2009, 327(2), 465-477.