JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Mutational and Structural Analysis of Human Brain Representing Intraspecific Variability

Anmol Juyal[1] Uma Kumari[2]

[1]Trainee, Rapture Biotech International Pvt. Ltd, Noida-201301, India, [2]Bioinformatics Scientist, Rapture Biotech International Pvt. Ltd, Noida-201301, India

***Abstract*:** The genomic landscape shows marked variation in the distribution of a number of features, including genes, transposable elements, GC content, CpG islands and recombination rate. Mutations were found through polymerase chain reaction-single strand conformation polymorphism, and then changed DNA fragments were purified and multiplied in sequencing. BLAST and ORF finder were used to find out the amino changes in sequence. The study is about 1FDQ, a Fatty-Acid Binding Protein (FABP). Taken from protein database, it shows crystal structure of a Brain Fatty-Acid Binding Protein (B-FABP) which is related to neuronal differentiation during embryonal neurodevelopment the absence of which has been shown to affect learning, memory, emotion regulation in rats – showing symptoms similar to schizophrenia. Similar is another protein called, Neurofilament M, a major component in the pathogenesis of Parkinson's disease (PD) where mutation analysis of the NF-M gene has been done in 322 familial PD patients. Two polymorphisms (Ala475Thr and Gly697Arg) occurring at similar frequencies in PD patients in substitutions affect residues of the NF-M protein that are highly conserved among different species. In analysis rare variants of the NF-M gene may act as susceptibility factors for PD and functional analyses of the identified variations are warranted to decipher possible mechanisms in neurodegeneration. Python-based software, alongside many Python plugin tools, has been developed to enhance the utilities and facilitate drug design in PyMOL which was used to visualize the protein structure, along with RasMol- another commonly-used visualization software. Various molecular modeling modules for visualization and analysis enhancement, protein–ligand modeling, molecular simulations, and drug screening are present in PyMOL. The human brain has been utilising cellular and molecular biological techniques for generations, where the mechanisms underlying the development, differentiation and function of the human brain remain quite elusive.

Index Terms- **Gene expression, human neuron protein, structural analysis, data compilation, fatty-acid binding protein.**

I. INTRODUCTION

Aeons of evolution has led to the creation of wonderfully complex biological structures with intricate mechanisms working in unison to create what we call life. One such example of this multigenerational phenomenon of trial and error is the human nervous system - more specifically, the brain. Human brain is an extremely dense mass of interconnected biological cellular wires called neurons which work like military estafettes- carrying signals originating from the brain and spinal cord in the form of electrical signals to stimulate different parts of the body in sequential manner for their routine functioning. Proteins play a big role in the functioning of our nervous system and, regarding the area of this study, are of two types - histone and non-histone proteins. Histone proteins are highly alkaline in nature and pack the DNA into small packages called nucleosomes which are the structural constituents of chromatin material. They are basically spools around which DNA wraps itself and are one of the most conservative in nature across all species of proteins. The electrostatic attractive forces between histone protein and the negatively-charged DNA causes the phosphate backbone to adhere to bind to the histone. Acetylation of lysine relieves the histones of some of its positive charge which increases repulsive forces leading to unwinding of DNA for transcription and thus playing a key role in gene expression. Histone proteins are of 5 types- H1(or H5), H2A, H2B, H3 and H4. These are divided further in 2 more classes - namely, core histones and linker histones. Dimers of H2A and H2B and tetramers of H3 and H4 form the octameric core of a nucleosome around which DNA wraps. Histones H1 and H5 are linkers located at the starting and end of the DNA sequence and attach it onto the nucleosome. Non-histone protein is what is left after the histones are removed. They

work like scaffolding to the DNA and work with histones to further condense the genetic material into chromosomes and other higher-level structures like chromatin material. Some examples of nonhistone proteins are DNA polymerase, polycomb, heterochromatin protein 1, scaffold proteins and other motor proteins. Their functions are wide-ranging - from DNA replication, transcription to nuclear transport, steroid hormone action, mitosis transition, etc. Classes of proteins that come into this category are- structural, regulatory and motor proteins. They are acidic and far less conservative in nature than histones as they are subject to more frequent occurrence of mutations - having a globally-aligned sequence in frequently encountered instances. Neuronal connections require the use of proteins and lipids for their structure as well as their metabolic activities. One such group of metabolically-important proteins is fatty-acid binding proteins (FABPs). They function as energy sources for a cell's metabolic functions, regulate metabolism in cell signalling processes, and membrane substrates (1. Furuhashi, M. et al., 2. McArthur, MJ. et al.). Fatty-acids, being insoluble in nature, makes them easy to be chaperoned to various cell organelles like nucleus, endoplasmic reticulum, mitochondria, peroxisomes and lipid droplets. Fatty-acid binding proteins show high specificity to these free ligands and bind to them. They are found mostly in organs closely related to fatty acid metabolism and have uses for them like liver, epidermis, brain, testes etc. The fatty-acid binding protein family comprises of 9 members coded by genes according to the area where they are found, these are adipocyte- (A-FABP), epidermal- (E-FABP), heart- (H-FABP), brain- (B-FABP), myelin- (M-FABP), intestine- (I-FABP), ileal-(Il-FABP), testis-FABP (T-FABP) and liver(L-FABP). The names merely suggest the predominant existence of these FABPs in these organs but does not signify their specificity towards that organ. The protein of interest of this study is a B-FABP called 1FDQ. Identified at chr 6q22-q23, the brain fatty-acid binding protein gene, also called FABP7, has been observed to show gene expression essential to brain development during mid-term embryonic stage in decreasing levels with progressional development (3. Feng, L. et al.). Genetic and regional expression by FABP7's regulatory components is observed in the human brain and central nervous system. Areas of CNS other than the brain include postnatal embryonic spinal cord, cerebellum and cerebral cortex (3. Feng,L.et al.). Experimentation with mice indicated prominent expression in the cerebellum at developing as well as postnatal stages (postnatal day 0-10). Gene expression decreased as the mice reached a "young adult" stage at day 20 which reduced to negligible amounts in healthy adults. Strong gene expression has been observed in developing brains, specifically in radial glial cells in the pre perinatal striatum which diminishes in mature white matter glia. Interesting to note is the fact that E-FABP and H-FABP are expressed in grey matter neurons as opposed to B-FABP (4. Haunerland, NH. et al.). B-FABP has high affinity for long chain fatty acids; with the preferred ligands being- n3 polyunsaturated fatty acids like α-linolenic acid, EPA (eicosapentaenoic acid) and DHA (docosahexaenoic acid). Affinities for n3 and n6 polyunsaturated FAs and monounsaturated n-9 FAs shown by titration calorimetric binding studies are Kd =27-53 nM, 115-206 nM and 41-47 nM, respectively (5. Balendiran, GK. et al.). It shows low binding affinities for saturated long chain fatty acids and no affinity for retinoic and palmitic acids. brain membrane phospholipids have highly concentrated amounts of polyunsaturated fatty acids EPA, DHA and arachidonic acid as compared to other tissues which may further shine a light to how B-FABP functions to ensure a constant supply of FA during embryonic cell maturation stage in prenatal brain as the CNS develops. Studies that involved reducing genetic expression in mice showed that B-FABPs affects memory, emotion, learning and brain development (6. Owada, Y. et al.). It is possible to have viable populations of mice lacking B-FABP with no morphological or histological abnormalities. Certain behavioural traits were exhibited that included enhanced anxiety, increased fear memory and response, increase in the amounts of arachidonic and palmitic acids in amygdala, decreased amount of NMDA receptor-mediated (N-methyl D aspartate) current to DHA and low levels of docosahexaenoic acid during prenatal period. An endophenotype involving deprecation prepulse inhibition was observed which is characteristic to schizophrenia (6. Owada, Y. et al., 7. Watanabe, A. et al.). Neurogenesis also got attenuated as a result of B-FABP absence. The migration of developing neurons into cortical layers is possible due to the influence of B-FABP - as suggested in a study by Feng, L. et al (3. Feng, L., et al.). B-FABP has been revealed to have high similarity to H-FABP (8. Bohmer, F.D. et al., 9. Specht, D. et al.), showing strong genetic expression in the mammary gland. When overexpressed, it inhibited tumour growth in a mouse breast cancer model (10. Hohoff, C. at al., 11. Maeda, K. et al.).

II. MATERIALS AND METHODS

Bioinformatics has been an indispensable tool in the construction, or even the conception of this paper. Protein databases by WWPDB and NCBI were employed to do studies on the protein sequence for type of alignment and its constituents, for protein comparison to other FABP sequences, physiochemical structure, especially the crystal structure of the protein. The Protein Data Bank was established by Brookhaven National Laboratory in 1969 and contains the largest openly-accessible database of 3D structures of macromolecules like proteins and nucleic acids. The database is curated and maintained by Worldwide Protein Data Bank and is in the service of biologists, geneticists and biochemists from around the world - making use of the vast repository of information as their routine lab practice as part of their research work. NMR, X-ray diffraction and electron microscopy has been used to determine approximate atomic coordinates, distance between pairs of atoms and protein conformation. Free, open-source software PyMOL and RasMol were employed for study and analysis of various physical and chemical aspects of 1FDQ protein. PyMOL is a computer software that produces high-quality 3-D images of micro and macromolecular biomolecules such as proteins. The "Py" in PyMOL refers to the computer program the software is written in which is Python. RasMol is another open-source program that was used to explore and visualise macromolecular structures in 3 dimensions in high quality. These computer programs have become indispensable tools in the field of structural biology research due to their practicality and accessibility. NCBI with its vast database of proteins, nucleotides, genetic sequences along with tools like BLAST, ClustalΩ and an online digital library of books and research papers is one of the cornerstones of bioinformatic tools to study biomolecules and biochemicals. ClustalΩ's MSA (Multiple Sequence Alignment was utilised to compare 1FDQ to other FABPs and analyse the type of sequence alignment, mutations and constituent elements and molecules that are involved in the structural and functional properties of the brain fatty-acid proteins. .Crystal structure of human brain (1FDQ) B-FABP has low binding affinity for saturated long chain fatty acids. Expression of brain fatty acid-binding protein (B-FABP) is correlated with neuronal differentiation during brain development. The bioinformatics tools mentioned above have been used to produce high-quality research as well as provide education in various institutions and laboratory facilities around the globe.

III. RESULTS AND DISCUSSION

1FDQ is a brain fatty acid-binding protein (B-FABP) which is spatially correlated with neuronal differentiation during brain development. The recombinant human B-FABP clearly exhibits high affinity for the polyunsaturated n-3 fatty acids alpha-linolenic acid, eicosapentaenoic acid, docosahexaenoic acid, and for monounsaturated n-9 oleic acid over polyunsaturated n-6 fatty acids, linoleic acid, and arachidonic acid (K(d) from 115 to 206 nm). B-FABP has low binding affinity for saturated long chain fatty acids where observing the three-dimensional structure of recombinant human B-FABP in complex with oleic acid shows "U-shaped" conformation. A comparison of the three-dimensional structures and binding properties of human B-FABP with other homologous FABPs, analysis of the primary and tertiary structures of human B-FABP provides a rationale for its high affinity and specificity for polyunsaturated fatty acids.1FDQ analysis in RasMol and PyMOL software to Crystal structure of human brain fatty acid binding proteinwere useful for study of a variety of macromolecules. Here the coloring is defined to identify the amino acid. Theouter polar part of protein is visible.

1FDQ structure is composed of 2218 atoms with a total molecular weight of 30.21kDa with262 being the total number of polymer monomer residues. With a sequence length of 131bp, the dual chain structure analysis indicates the existence of HXA coding for Docosa-4,7,10,13,16,19-hexaenoic acid (DHA) – a B-FABP ligand which is common feature observed across the FABP family at the A and B chain terminals at positions 131 and 631 respectively with a binding affinity of Kd**:** 53.4nM. Of the total number of 131 residues, the element makeup of the two chains is C: 653, N: 175, O: 204, S: 5.

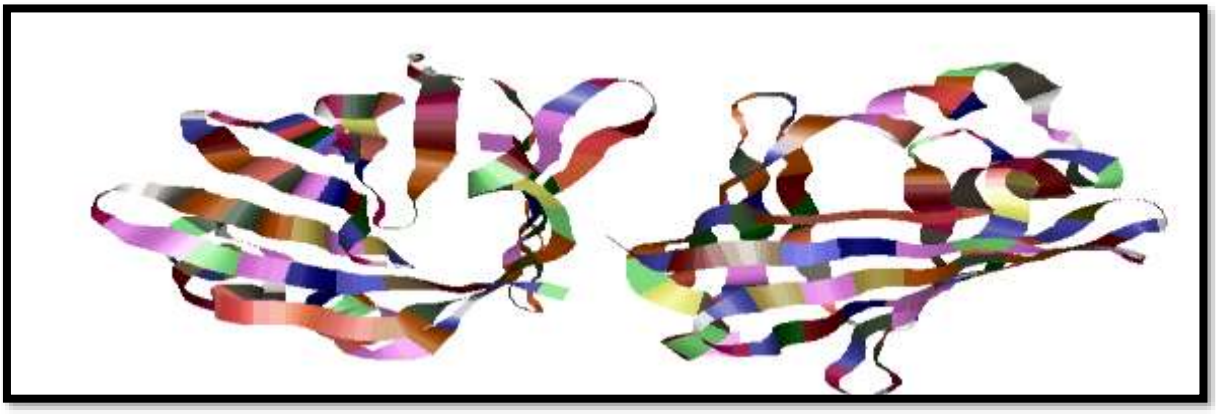

Figure 1. "Shapely" colour scheme colour codes residue by amino acid property.

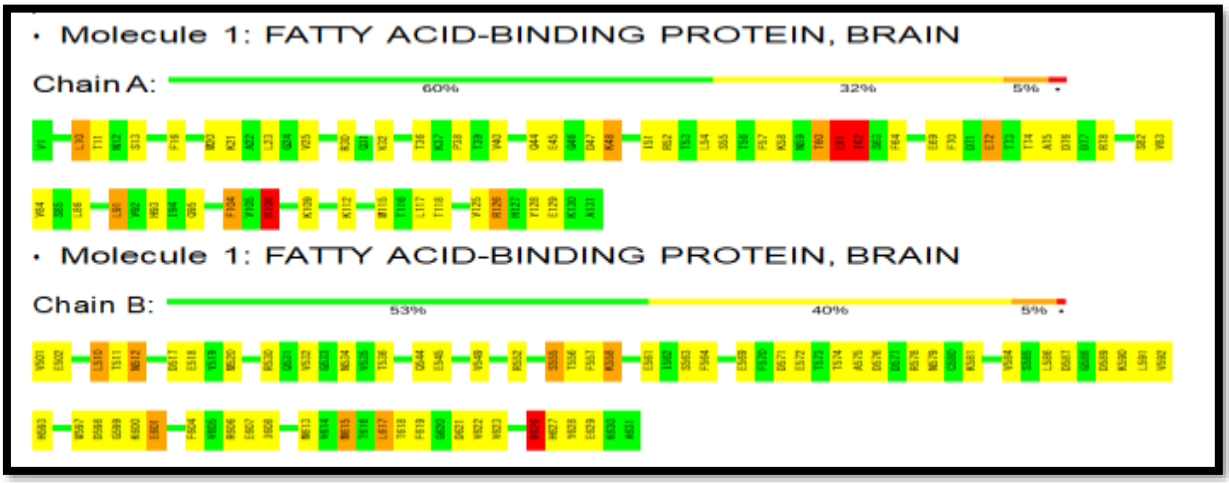

Figure 2. These plots are drawn for all protein, RNA and DNA chains in the entry. The graphic for A chain summarises the proportions of the various outlier classes displayed in the second graphic (chain B). This graphic shows the sequence view annotated by issues in geometry. Residues are colour-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

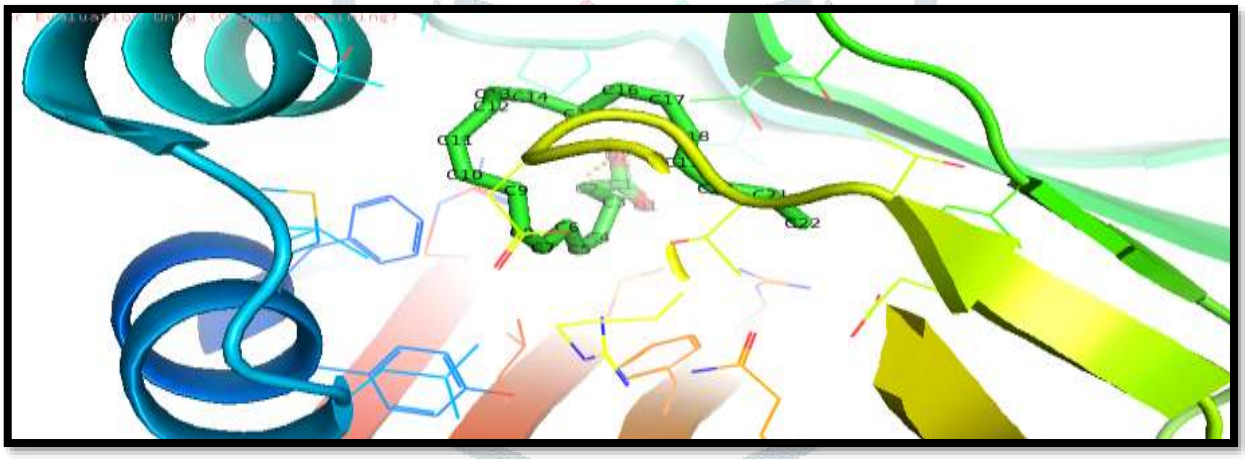

Figure 3. HXA Ligand visualisation in PyMOL

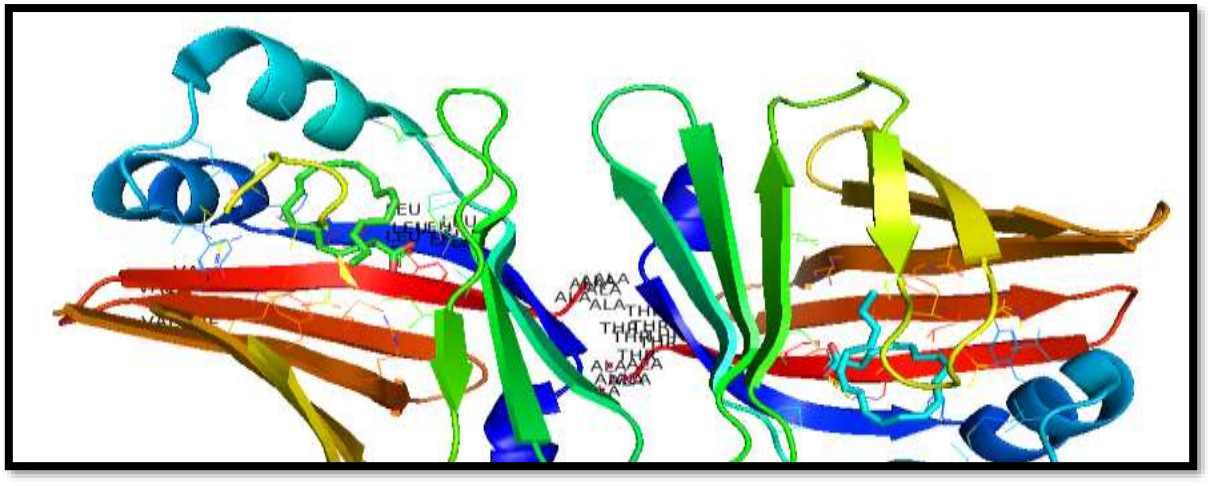

Figure 4. N terminal and C terminal visualization in PyMOL (sequence starts at the N-terminus (Ala)

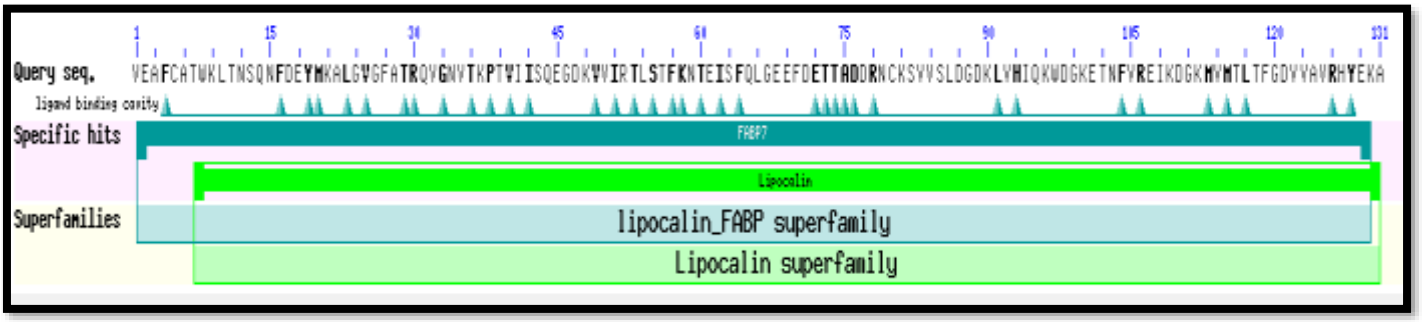

Figure 5. Fatty acid binding protein 7; FABP7 (also known as brain FABP, B-FABP, BLBP, brain lipid binding protein) is highly expressed in glial cells through development of the nervous system. In the developing brain, FABP7 is required for the establishment of the radial glial fiber system - involved in the migration of immature neurons. This subgroup belongs to the intracellular fatty-acid binding protein (FABP) family, members of which are small proteins that bind hydrophobic ligands in a non-covalent, reversible manner, and have been implicated in intracellular uptake, transport and storage of hydrophobic ligands, regulation of lipid metabolism and sequestration of excess toxic fatty acids, as well as in signaling, gene expression, inflammation, cell growth and proliferation, and cancer development.

IV. CONCLUSION

The changes occurred in the protein structure caused by point mutation may be the reason for tumorigenesis of astrocytoma. In RasMol the amino colour scheme shows the amino acids according to traditional amino acid properties. The colours are there to identify amino acids. In the outer part of the protein, polar regions are visible in bright colors and non-polar residues darker. In colour, the "Shapely" scheme which can be selected from the RasMol Colours menu, ASP, GLU are bright red and CYS, MET represented in yellow. The 3D visualization color-codes each atom according to the anisotropic temperature (B-factor) value stored in the PDB file. The difference between the temperature and charge color schemes is that increasing temperature values proceed from blue to red, whereas increasing charge values go from red to blue. The RasMol "group" properties of color codes residues identified by their position in a macromolecular chain. Each chain is drawn as a smooth spectrum from blue through green and yellow and orange to red. Structural analysis of N terminal of proteins are colored blue (similar to the CPK color for nitrogen) and the C termini, red (similar to the CPK color for oxygen). Conserved domain analysis of the Lipocalin / cytosolic fatty-acid binding protein family shows that they act as transporters for small hydrophobic molecules, such as lipids, steroid hormones, bilins, and retinoids. The sequence alignment subsumes both the lipocalin and fatty acid binding protein signatures from PROSITE, supported structurally and functionally. Where the structure is an eight-stranded beta barrel.

REFERENCES

[1] Furuhashi M, Hotamisligil GS: Fatty acid-binding proteins: Role in metabolic diseases and potential as drug targets. Nat Rev Drug Discov. 2008, 7: 489-503. 10.1038/nrd2589.

[2] McArthur MJ, Atshaves BP, Frolov A, Foxworth WD, et al: Cellular uptake and intracellular trafficking of long chain fatty acids. J Lipid Res. 1999, 40: 1371-1383.

[3] Feng, L., Hatten, M. E. & Heintz, N. Brain lipid-binding protein (BLBP): a novel signaling system in the developing mammalian CNS. Neuron 12, 895–908 (1994).

[4] Haunerland NH, Spencer F: Fatty acid-binding proteins: Insights from genetic manipulations. Prog Lipid Res. 2004, 43: 328-349. 10.1016/j.plipres.2004.05.001.

[5] Balendiran GK, Schnutgen F, Scapin G, Borchers T, et al: Crystal structure and thermodynamic analysis of human brain fatty acid-binding protein. J Biol Chem. 2000, 275: 27045-27054.

[6] Owada, Y., Abdelwahab, SA., Kitanaka, N., Sakagami, H., et al: Altered emotional behavioural responses in mice lacking brain-type fatty acid-binding protein gene. Eur J Neurosci. 2006, 24: 175-187. 10.1111/j.1460-9568.2006.04855.

[7] Watanabe, A. et al. Fabp7 maps to a quantitative trait locus for a schizophrenia endophenotype. PLoS Biol.5, 2469–2483 (2007).

[8] Bohmer, F. D. et al. Identification of a polypeptide growth inhibitor from bovine mammary gland. Sequence homology to fatty acid- and retinoid binding proteins. J. Biol. Chem. 262, 15137–15143 (1987).

[9] Specht, B. et al. Mammary derived growth inhibitor is not a distinct protein but a mix of heart-type and adipocyte-type fatty acid-binding protein. J. Biol. Chem. 271, 19943–19949 (1996).

[10] Hohoff, C. & Spencer, F. Correspondence re: Y. E. Shi et al., Anti tumour activity of the novel human breast cancer growth inhibitor, mammary-derived growth inhibitor-related gene, MRG. Cancer Res. 57, 3084–3091, 1997. Cancer Res. 58, 4015–4017 (1998).

[11] Maeda, K. et al. Adipocyte/macrophage fatty acid binding proteins control integrated metabolic responses in obesity and diabetes. Cell. Metab. 1, 107–119 (2005).

[12] Anderson, S. Shotgun DNA sequencing using cloned DNase I-generated fragments. *Nucleic Acids Res.* **9**, 3015–3027 (1981).

[13] Gardner, R. C. et al. The complete nucleotide sequence of an infectious clone of cauliflower mosaic virus by M13mp7 shotgun sequencing. *Nucleic Acids Res.* **9**, 2871–2888 (1981).

[14] Venter, J. C. et al. Shotgun sequencing of the human genome. *Science* 280, 1540–1542 (1998).

[15] Lee, L. G. et al. New energy transfer dyes for DNA sequencing. *Nucleic Acids Res.* 25, 2816–2822 (1997).

[16] Mahfouz A, van de Giessen M, van der Maaten L, Huisman S, Reinders M, Hawrylycz MJ, et al. Visualizing the spatial gene expression organization in the brain through non-linear similarity embeddings. Methods. 2015; 73:79–89. Epub 2014/12/03. pmid:25449901.

[17] Marchler-Bauer A et al. (2017), *"CDD/SPARCLE: functional classification of proteins via subfamily domain architectures."*, Nucleic Acids Res.45(D)200-3.