

The GTOP database in 2009: updated content and novel features to expand and deepen insights into protein structures and functions

Satoshi Fukuchi^{1,*}, Keiichi Homma¹, Shigetaka Sakamoto², Hideaki Sugawara¹, Yoshio Tateno¹, Takashi Gojobori¹ and Ken Nishikawa³

¹Center for Information Biology and DNA Data Bank of Japan, National Institute of Genetics, Yata 1111, Mishima, Shizuoka 411-8540, ²HOLONICS Corporation, Soeji 85, Numazu, Shizuoka 411-0803 and ³Department of Bioinformatics, Maebashi Institute of Technology, Kamisadori 460-1, Maebashi, Gunma 371-0816, Japan

Received September 12, 2008; Revised October 15, 2008; Accepted October 16, 2008

ABSTRACT

The Genomes TO Protein Structures and Functions (GTOP) database (<http://spock.genes.nig.ac.jp/~genome/gtop.html>) freely provides an extensive collection of information on protein structures and functions obtained by application of various computational tools to the amino acid sequences of entirely sequenced genomes. GTOP contains annotations of 3D structures, protein families, functions, and other useful data of a protein of interest in user-friendly ways to give a deep insight into the protein structure. From the initial 1999 version, GTOP has been continually updated to reap the fruits of genome projects and augmented to supply novel information, in particular intrinsically disordered regions. As intrinsically disordered regions constitute a considerable fraction of proteins and often play crucial roles especially in eukaryotes, their assignments give important additional clues to the functionality of proteins. Additionally, we have incorporated the following features into GTOP: a platform independent structural viewer, results of HMM searches against SCOP and Pfam, secondary structure predictions, color display of exon boundaries in eukaryotic proteins, assignments of gene ontology terms, search tools, and master files.

INTRODUCTION

Proteins encoded by genomes generally function after adopting proper 3D structures. A rapid increase in the number of entirely sequenced genomes led to an unprecedented growth in the number of hypothetical proteins

resulting from genome annotation. Protein structures and functions can be inferred from amino acid sequences by using advanced computer programs. There is no doubt in the importance of structural and functional annotations of hypothetical proteins. The GTOP project was started in 1999 as reported (1) and was taken over by the DNA Data Bank of Japan (2) in 2007, under which the database has been continuously updated. GTOP is a database that provides protein annotation of 3D structures and functions based on similarity searches against PDB (3), SCOP (4), and Swiss-Prot (5), 2D structure predictions, Pfam (6) protein families, PROSITE (7) functional motifs, prediction of trans-membrane regions, and others.

There are several databases of the 3D structures of all the genome-encoded proteins. For example, SUPER-FAMILY (<http://supfam.mrc-lmb.cam.ac.uk/SUPER-FAMILY/>) (8) provides SCOP domain assignments to proteins encoded by completely sequenced genomes. A collection of comparative protein 3D structure models is available at Modbase (<http://modbase.compbio.ucsf.edu/modbase-cgi/index.cgi>) (9) in some entirely sequenced genomes. Gene3D (<http://gene3d.biochem.ucl.ac.uk/Gene3D/>) (10) makes public CATH-based domain assignments and functional annotations to proteins in more than 500 genomes. Functional and domain assignments including intrinsically disordered (ID) regions can be found at PEDANT (<http://pedant.gsf.de/>) (11).

From the previous report, we have added a large body of data and tools to GTOP, for example ID region assignments, exon information on eukaryotic proteins, an efficient mechanism to search within a user-specified set of genomes, and tools for phylogenetic profile search. Since its inception, GTOP has employed a user-friendly interface to let the user grasp features of a query protein at a glance. The interface has been improved with the addition of new information. A GTOP user can readily obtain

*To whom correspondence should be addressed. Tel: +81 55 981 6837; Fax: +81 55 981 6889; Email: sfukuchi@genes.nig.ac.jp

© 2008 The Author(s)

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/2.0/uk/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

comprehensive structural and functional data of all the proteins encoded by entirely sequenced genomes.

UPDATE IN GTOP THAT CONTRIBUTED TO IMPROVED STRUCTURAL ASSIGNMENTS

A list of the genomes stored in GTOP is available at <http://spock.genes.nig.ac.jp/~genome/org.html>, together with the abbreviations of organism names used in the database. In the 2002 paper, we reported that GTOP contained protein data of 41 genomes (1). The database has grown to cover a total of 797 genomes, with 41, 466, 114 and 176 genomes of archaea, eubacteria, eukaryota and bacteriophages, respectively. The following data are subject to regular renewal: (i) amino acid sequences encoded by genomes newly sequenced after the previous update, (ii) amino acid sequences that existed in the previous version but were subsequently modified and (iii) reference databases such as PDB, SCOP, Swiss-Prot, Prosite, and Pfam whose new versions were released. The sequences fallen in category (ii) were recalculated to keep annotations up-to-date. Update category (iii) is crucial to keep annotations up-to-date, because most annotations in GTOP are obtained by homology search programs or those based on homology search.

The main focus of GTOP is structural annotations made by homology searches against the PDB and SCOP databases. Although GTOP used PSI-BLAST (12) in the previous report, it now employs reverse-PSI-BLAST (13), as this method gives comparable results in drastically reduced computation time. HMM searches using the SUPER-FAMILY profiles (8) of SCOP domains were additionally conducted, as they are particularly effective in identifying small domains such as DNA binding domains.

Figure 1 presents a time course of the number of the genomes stored and the average fractions of proteins with

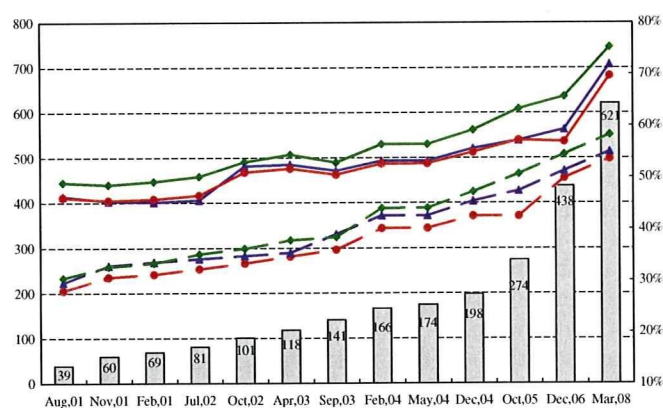


Figure 1. The time courses of the number of genomes included and the fraction of the sequences with homologs in the PDB. The line graphs represent the ratios of the sequences with homologs in the PDB, while the column graph stands for the number of genomes in GTOP. The scales for the fraction and the number of genomes are shown at the right and left ends, respectively. The blue, green, and red lines correspond to fruit fly, *E. coli*, and the overall average, respectively. The solid and dotted lines respectively show the ratios obtained using reverse-PSI-BLAST, and those using BLAST. The exact numbers of genomes are displayed near the top of the rectangles.

3D annotations made by BLAST and reverse-PSI-BLAST. The fraction of sequences with alignments to PDB shows a steadily increasing trend, reflecting the growth of the PDB database. The fraction aligned by reverse-PSI-BLAST exceeds that by BLAST, reflecting the higher sensitivity of the former method. However, one should note that in this statistics a sequence is considered to be annotated if it has at least one PDB hit by BLAST or reverse-PSI-BLAST and it may have large tracts of structurally undetermined regions. When statistics is evaluated residue-wise, the fractions of regions aligned to PDB sequences in the latest version in human and *Escherichia coli* proteins are 47% and 64%, respectively.

ID REGIONS

As most proteins do not entirely consist of structural domains, the fraction of residues with structural assignments will not reach unity; outside of globular domains there exist ID regions that assume no specific 3D structures by themselves, and tend to contain active regions in proteins involved in crucial biological processes such as signal transduction and transcriptional regulation (14–16). Recent research revealed that ID regions exist predominantly on the cytoplasmic side of eukaryotic proteins (17), play important roles in cell signaling, transcriptional control (18). We predicted ID regions in proteins stored in GTOP by the DISOPRED2 (19) program and presented them. Figure 2A shows a GTOP screen shot of human androgen receptor, a typical protein with long ID regions. As this example illustrates, GTOP graphically displays complex domain architectures of eukaryotic proteins composed of structural domains and ID regions.

EXON BOUNDARIES IN EUKARYOTIC PROTEINS

The existence of introns and exons is a unique feature of eukaryotic genes and the location of exon boundaries in the corresponding protein structure is of interest (20). We thus developed tools to display exon boundaries on amino acid sequences and 3D structures. Figure 2B shows an example of the exon boundary view. The exons are presented in 5 colors both in the 3D structure and the sequence displays, from which the boundaries can be clearly seen. We developed a 3D viewing system incorporating Jmol applet (<http://www.jmol.org/>) so that the user can view 3D structures in the browser without installing additional software. Alternatively Rasmol (21) or Chime (<http://www.mdl.com/>) can be used. Exon information is also presented in green and blue stripes (near the bottom of Figure 2A).

SEARCH TOOLS

GTOP strives to keep precomputed annotations of all the amino acid sequences of proteins derived from all the completely sequenced genomes. One clear benefit of having precomputed annotations beside the rapidity of supplying information is to make inter-genomic

comparative analyses possible. Phylogenetic profile search is one analytical tool that exploits this advantage: a user-specified search produces the presence and absence pattern of features such as SCOP folds, superfamilies, and families, Pfam domains, PROSITE motifs, and the number of trans-membrane helices. The user can conduct a search for a specific feature that are present in certain species and/or absent in others; for example, a search for a SCOP domain present in all the eubacterial species and absent in all the eukaryotic species in GTOP. The summary section of GTOP also offers comparative statistics, which has the ratio of 3D annotations in each genome, the frequencies of SCOP folds, superfamilies, and families, Pfam domains and PROSITE motifs.

Expansion of the database resulted in increased search time. The tools for keyword, homology, and text searches in GTOP were thus modified so that the user can reduce search time through selection of the genomes in which to conduct a search. The user can easily specify organisms with the use of check boxes placed next to organism names.

MASTER FILES

An annotation summary of each protein, consisting of abbreviated one-line descriptions, is saved in a master file. Master file information for each protein is displayed below a GTOP diagram of the type shown in Figure 2A. All the available data of each genome have been compiled in one file, freely downloadable from <ftp://spock.genes.nig.ac.jp/pub/gtop/>. Explanations of the meanings for each HEADER can be found at <http://spock.genes.nig.ac.jp/~genome/mas-doc.html>.

FUTURE DIRECTIONS

Despite the wealth of currently available structural data and use of sensitive programs, considerable fractions of most proteins have neither structural domains nor ID regions assigned. We are currently developing a system to accurately classify the fraction into structural domains and ID regions. Excitingly this will result in reliable identification of structural domains whose 3D structures remain undetermined. We expect that the installation of this system will provide further insights into the protein structure. We are also considering incorporation of protein-protein interaction data to enrich GTOP further.

FUNDING

The GTOP database is supported in part by the Target Protein Research Program from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and in part by the Bioinformatics Research and Development Project from the Japan Science and Technology Agency. Funding for open access publication charge: the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Conflict of Interest statement: None declared.

REFERENCES

- Kawabata, T., Fukuchi, S., Homma, K., Ota, M., Araki, J., Ito, T., Ichiyoshi, N. and Nishikawa, K. (2002) GTOP: a database of protein structures predicted from genome sequences. *Nucleic Acids Res.*, **30**, 294–298.
- Sugawara, H., Ogasawara, O., Okubo, K., Gojobori, T. and Tateno, Y. (2008) DDBJ with new system and face. *Nucleic Acids Res.*, **36**, D22–D24.
- Henrick, K., Feng, Z., Bluhm, W.F., Dimitropoulos, D., Doreleijers, J.F., Dutta, S., Flippen-Anderson, J.L., Ionides, J., Kamada, C., Krissinel, E. *et al.* (2008) Remediation of the protein data bank archive. *Nucleic Acids Res.*, **36**, D426–D433.
- Andreeva, A., Howorth, D., Chandonia, J.M., Brenner, S.E., Hubbard, T.J., Chothia, C. and Murzin, A.G. (2008) Data growth and its impact on the SCOP database: new developments. *Nucleic Acids Res.*, **36**, D419–D425.
- UniProt Consortium. (2008) The universal protein resource (UniProt). *Nucleic Acids Res.*, **36**, D190–D195.
- Finn, R.D., Tate, J., Misty, J., Coghill, P.C., Sammut, S.J., Hotz, H.R., Ceric, G., Forslund, K., Eddy, S.R., Sonnhammer, E.L. *et al.* (2008) The Pfam protein families database. *Nucleic Acids Res.*, **36**, D281–D288.
- Hulo, N., Bairoch, A., Bulliard, V., Cerutti, L., Cuche, B.A., de Castro, E., Lachaize, C., Langendijk-Genevaux, P.S. and Sigrist, C.J. (2008) The 20 years of PROSITE. *Nucleic Acids Res.*, **36**, D245–D249.
- Wilson, D., Madera, M., Vogel, C., Chothia, C. and Gough, J. (2007) The SUPERFAMILY database in 2007: families and functions. *Nucleic Acids Res.*, **35**, D308–D313.
- Pieper, U., Eswar, N., Davis, F.P., Braberg, H., Madhusudhan, M.S., Rossi, A., Marti-Renom, M., Karchin, R., Webb, B.M., Eramian, D. *et al.* (2006) MODBASE: a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Res.*, **34**, D291–D295.
- Yeats, C., Lees, J., Reid, A., Kellam, P., Martin, N., Liu, X. and Orengo, C. (2008) Gene3D: comprehensive structural and functional annotation of genomes. *Nucleic Acids Res.*, **36**, D414–D418.
- Riley, M.L., Schmidt, T., Artamonova, I.I., Wagner, C., Volz, A., Heumann, K., Mewes, H.W. and Frishman, D. (2007) PEDANT genome database: 10 years online. *Nucleic Acids Res.*, **35**, D354–D357.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
- Marchler-Bauer, A., Panchenko, A.R., Shoemaker, B.A., Thiessen, P.A., Geer, L.Y. and Bryant, S.H. (2002) CDD: a database of conserved domain alignments with links to domain three-dimensional structure. *Nucleic Acids Res.*, **30**, 281–283.
- Dunker, A.K., Lawson, J.D., Brown, C.J., Williams, R.M., Romero, P., Oh, J.S., Oldfield, C.J., Campen, A.M., Ratliff, C.M., Higgs, K.W. *et al.* (2001) Intrinsically disordered protein. *J. Mol. Graph. Model.*, **19**, 26–59.
- Tomba, P. (2005) The interplay between structure and function in intrinsically unstructured proteins. *FEBS Lett.*, **579**, 3346–3354.
- Wright, P.E. and Dyson, H.J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.*, **293**, 321–331.
- Minezaki, Y., Homma, K. and Nishikawa, K. (2007) Intrinsically disordered regions of human plasma membrane proteins preferentially occur in the cytoplasmic segment. *J. Mol. Biol.*, **368**, 902–913.
- Minezaki, Y., Homma, K., Kinjo, A.R. and Nishikawa, K. (2006) Human transcription factors contain a high fraction of intrinsically disordered regions essential for transcriptional regulation. *J. Mol. Biol.*, **359**, 1137–1149.
- Ward, J.J., Sodhi, J.S., McGuffin, L.J., Buxton, B.F. and Jones, D.T. (2004) Prediction and functional analysis of native disorder in

- proteins from the three kingdoms of life. *J. Mol. Biol.*, **337**, 635–645.
20. Homma,K., Kikuno,R.F., Nagase,T., Ohara,O. and Nishikawa,K. (2004) Alternative splice variants encoding unstable protein domains exist in the human brain. *J. Mol. Biol.*, **343**, 1207–1220.
21. Sayle,R.A. and Milner-White,E.J. (1995) RASMOL: biomolecular graphics for all. *Trends Biochem. Sci.*, **20**, 374.
22. Kumar,R., Betney,R., Li,J., Thompson,E.B. and McEwan,I.J. (2004) Induced alpha-helix structure in AF1 of the androgen receptor upon binding transcription factor TFIIF. *Biochemistry*, **43**, 3008–3013.

V. 參考資料

参考資料に関する一覧表

掲載紙名	掲載日	見出し
Medical Tribune	2010年2月25日	PEG-IFN α +RBV 併用療法の有効性を規定する SNPs を同定
朝日新聞 (東京版)	2009年9月14日	C型肝炎治療の効き目、患者遺伝子でも差、日本の研究班
朝日新聞	2009年9月14日	C型肝炎治療効果 DNAで差
日本経済新聞	2009年9月14日	C型肝炎治療、薬剤の効果予測、厚生省研究班 遺伝子の個人差解析
中日新聞	2009年9月14日	C型肝炎治療遺伝子が左右、効かない人は型に違い
東京新聞	2009年9月14日	C型肝炎治療の効果 遺伝子型で差、国内研究班



PEG-IFN- α + RBV併用療法の有効性を規定するSNPsを同定

C型慢性肝炎治療で薬剤投与前の判別が可能に

C型慢性肝炎に対するペグインターフェロン(PEG-IFN) α とリバビリン(RBV)の併用療法は、現時点で唯一根治が期待できる治療法である。しかし、半数の患者はゴールドスタンダードであるこの治療法でも有効性が得られない。有効例と無効例をなんとか薬剤投与前に判定できないか。名古屋市立大学大学院病原医学ウイルス分野の田中靖人教授は、国立国際医療センター国府台病院肝炎・免疫研究センターの瀧上雅史センター長、東京大学大学院人類遺伝学分野の徳永壽士教授らとともに、厚生労働省の班研究の一環としてC型慢性肝炎治療の有効性を規定する一塩基多型(SNPs)の解析に取り組み、危険率30倍ものSNPs同定に成功した。



田中靖人氏

治療効果を規定する 遺伝子多型検定の初め

C型肝炎は、感染者の30%が自然に治癒することが知られている。しかし残りの70%は慢性肝炎に移行し、肝硬変から肝がん発症に至る。現時点でこの不可逆的な進展を阻止できる唯一の手段は、PEG-IFN α とRBVの併用療法(以下PEG-IFN- α /RBV)である。しかし、C型肝炎ウイルス(HCV)のgenotypeが1型で、かつウイルス量も多いタイプに対しては、PEG-IFN- α /RBVは患者の5割でしか有効性が得られないことが明らかになっている。しかも、このタイプはわが国のC型慢性肝炎患者の7割を占める。

さらにPEG-IFN- α /RBVに対する反応性として、①ウイルス量が減りやすく、したがっておろしやすい人 ②ウイルス量が減りにくく、一時病状が改善しても再燃する可能性がある人 ③PEG-IFN- α /RBVが全く効かない人への3群に分かれるが、どの患者がどのタイプであるかは、実際に薬剤を3か月ほど継続投与してみなければ判定できない。

このような治療反応性を規定する因子としては、先述のHCVのgenotypeやウイルス量のほか、血小板数、線維化レベル、高齢女性であるなど、これまでにさまざまな因子の関与が指摘されてきた。あるいはウイルス遺伝子のコア領域やNSSAと呼ばれる非翻訳領域に変異があるとIFNが効きにくいとも言われている。

しかし、どの因子も治療反応性を規定する因子として決定的ではなく、かつ宿主側の遺伝要因についての検討は従来ほとんど行われていなかった。薬剤反応性にかかわる包括的なSNPsの検討は、今回が初の試み。

マイナーアリの 危険率30倍のSNPを検出

田中教授らは、PEG-IFN- α /RBVの治療効果と関連する宿主遺伝子を同定する目的で、全国19施設からC型慢性肝炎患者142例から検体の供与を受けた。患者の内訳は、ウイルス学的見地からPEG-IFN- α /RBV治療に反応しなかったnull virologic re-

sponse(NVR)群が78例、治療に反応したvirologic response(VR)群が64例だった。VR群には、持続的に良好な反応を示したsustained virologic response(SVR)群と、治療によって一過性にウイルス量が減少したものの再燃の可能性があるtransient virologic response(TVR)群が含まれる。

全例を対象にまずゲノムワイド関連解析を実施し、一挙に90万個のSNPsを検出。その結果、突出して高い危険率を示す2つのSNPsが染色体19番で同定された。すなわち、その点変異(マイナーアリ)を持つ場合、点変異を持たない(メジャーアリ)患者に比べて、PEG-IFN- α /RBV治療が無効である確率は、一方のSNPはP値 1.93×10^{-13} 、オッズ比(OR)20.3、もう一方のSNPはP値 3.11×10^{-15} 、OR 30.0と、20倍から30倍という高い危険率を持つことが明らかになった。

次に、ゲノムワイド関連解析の結果が正しいかどうかを確認するため、別の患者コホートを対象に行ったreplication解析でも、2つのSNPsの一方はP値 5.46×10^{-15} 、OR 19.2、もう一方はP値 9.47×10^{-16} 、OR 27.4で、ゲノムワイド関連解析とほぼ同等の結果を示し、同解析の正しさが証明された。

検討の全過程を通じて、解析対象となったC型慢性肝炎患者は314例で、内訳はNVR 128例、VR 186例で、VR中SVRが140例を占めた。またマイナー SNPをホモ接合体として持つ例は、NVR群のみに見られた。この結果は、Nature Geneticsの2009年9月13日付けオンライン版で報じられた。

マイナーアリを持つ患者は IL-28遺伝子の発現レベルが低い

ゲノムワイド関連解析で検出された2つのSNPsが存在する染色体19番には、IFN- λ 3とも呼ばれるインターロイキン(IL)-28B遺伝子がコードされている。その近傍には、同じくIFN- λ のサブタイプであるIL-28A(IFN- λ 2)遺伝子とIL-29(IFN- λ 1)遺伝子が存在するが、今回のSNPsはIL-28B遺伝子のみに関連し、PEG-IFN- α /RBVが効かなかった患者は

IL-28B遺伝子の発現レベルが有意に低い(P=0.015)ことも明らかになった。

これまでに報告されたSNPsはすべて世界共通の登録番号に相当する番号で表示されており、今回ゲノムワイド関連解析でOR 30だったSNPの番号はrs8099917である。メジャーアリがチロシン(T)であるのに対してグリシン(G)アリルを持つ。田中教授らの発表と同時期に米国のグループが同定したSNPはrs12979860で、やはり19番染色体のIL-28B遺伝子の近傍にあり、rs8099917と連鎖不平衡であることが判明した(図)。また、rs8099917の近傍を精査した結果、IL-28B遺伝子を含む15.7キロベースの範囲内に、rs8099917と連鎖不平衡の関係にあるSNPsは少なくとも6個存在し、ハプロタイプ解析からマイナーアリ(リスクアリル)を持つ場合の治療無効となるORは11.1(P= 1.35×10^{-25})であった。

さらにロジスティック回帰モデルによって、治療反応性について従来

指摘されてきた年齢、性、血小板数、線維化ステージ、ウイルス量などを含め、どの因子がNVRと最も関連するかを検討したところ、rs8099917と性(女性)の2つの因子のみが有意だった(表)。

同教授は今回の結果について、「OR 30は過去に例がないような大きな数値で、血液検査でSNPsを調べるだけで、現行の治療が効く人と効かない人を明確に判定することができる。検査項目として7個のSNPsのどれを使ってもよいが、IL-28B遺伝子のイントロン中のものでエクソン中にあるもの、および上流にあるものの3個を組み合わせたのがよいのではないかと考えている。7個のSNPsのどれが最も重要かは引き続き検討中だ」と話す。

PEG-IFN- α /RBV試験投与前に 回避できる

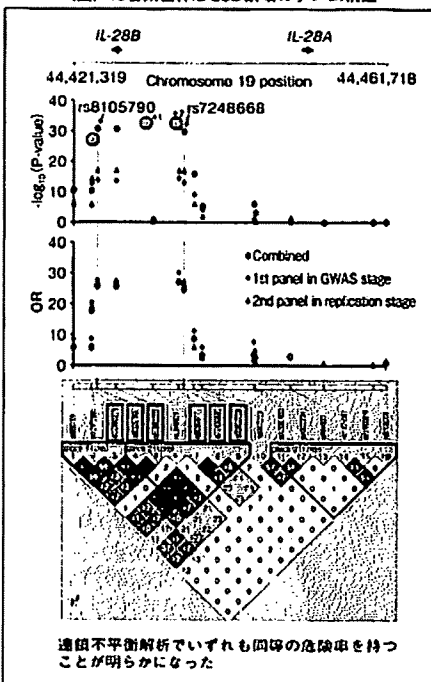
現行の3か月間の試験治療をする前にマイナーアリを持つか、メ

ジャーアリルを持つかを調べ、無効と判定されれば、患者は種々の副作用を回避できるだけでなく、国庫負担の軽減にも寄与することができる。

IFN- λ 3はIFN- α 、 β 、 γ と異なるレセプターを用いて抗ウイルス効果を発揮することが知られているが、レセプター結合後のシグナル伝達経路は重複する部分が多いため、IFN- α とIFN- λ 間の相互作用の詳細はまだ明らかにされておらず、その解明は田中教授らの今後のテーマの1つだが、IL-28B遺伝子の発現が低い場合、IFN- α のシグナル伝達経路がなんらかの形で阻害される可能性が考えられるという。

PEG-IFN- α /RBV無効例では、プロテアーゼ阻害薬を上乗せした3剤併用による著効(ウイルス排除)はそれほど大きく期待できないため、新たな治療法の確立が必要となる。今回の研究成果によってIFN- λ 製剤の登場も期待され、近い将来、C型慢性肝炎治療に新たな局面がもたらされる可能性が高い。

(図) 19番染色体IL-28B領域のゲノム構造



*: rs12979860(Go, et al.), **: rs8099917(Tanaka, et al.) (田中靖人氏提供)

(表) PEG-IFN- α /RBV治療無効(NVR)と関連する因子(ロジスティック回帰モデルによる)

因子	OR	95%信頼区間	P値
SNP/rs8099917 (Gアリル)	37.68	16.71~83.85	<0.0001
年齢	1.02	0.98~1.07	0.292
性(女性)	3.32	1.49~7.39	0.003
再治療	1.12	0.55~2.33	0.750
血小板数	0.93	0.87~1.01	0.080
ALTレベル	1.00	0.99~1.00	0.735
線維化ステージ	1.10	0.73~1.66	0.658
HCV-RNAレベル	1.01	0.99~1.02	0.139

(Nature Genetics 2009; 41 (10): 1105-1109より改変)

C型肝炎の治療が効くか効かないかは、人の遺伝子のわずかな違いが要因の一つになっていることが、国立国際医療センターの溝上雅史肝炎・免疫研究センター長と名古屋市立大の田中靖人准教授らのグループの研究でわかった。14日付の米科学誌ネイチャー・ジエネティクス(電子版)に発表される。

C型肝炎はウイルスが原因の病気で、日本人に最も多いタイプでは治療薬「インタフェロン」と抗ウイルス剤「リバビ

リン」の併用療法が有効とされている。しかし、約20%の患者は効きづらく、治療を受けてみないと効果がわからなかった。

このため、効きづらい患者は、月数万円の薬代が無駄になったり、発熱やうつ病などの副作用が出かねない危険を抱えながら治療を続けなければならなかった。近く数千円の遺伝子検査キットが開発できるといい、治療前に血液を検査することにより、こうした問題を避けられるという。

C型肝炎治療効果 DNAで差

C型肝炎治療

薬剤の効果予測

厚生労働省
研究班

遺伝子の個人差解析

厚生労働省研究班(班

長 田中靖人・名古屋市

立大学准教授)はC型肝炎

の一般的な治療法が効

きにくい患者を、遺伝子

の個人差から事前に見つ

けることのできる手法を

発見した。特定の遺伝子

配列を持つ患者は、通常

の患者と比べて併用療法

が効かない可能性が30倍

高まる。85〜90%の的中

率で治療法の効果を事前

に予測できるという。

米科学誌ネイチャー・

ジエネティクス(電子版)

に14日、発表する。

C型肝炎はウイルスの

感染で起こり、肝臓がん

の原因となる。C型肝炎

ウイルス感染者は推定で

約200万人いる。最も

効果の期待される「ペグ

インターフェロン」と「リ

バビリン」という薬剤の

併用療法は患者の約半数

にしか効かない。

国立国際医療センター

国府台病院の溝上雅史・

肝炎・免疫研究センター

長と名古屋市大の田中准

教授らは、併用療法の有

効いた患者186人と無効

だった128人の遺伝子

を解析。「IL28B」と

いう遺伝子の配列のわず

かな個人差によって治療

法の効果が大きく左右さ

れることを発見した。

事前に血液を採取して

遺伝子の個人差を調べる

だけで、併用療法の効果

が期待できるかを高い確率で予測できる。研究班は今後、臨床研究に乗り出す。この遺伝子を標的とした新たな治療薬の開発につながる可能性もあるという。

特定の遺伝子の型を
持つ人には、C型肝炎
の治療効果が山ないこ
とを、名古屋市立大の
甲中靖人准教授、国立
国際医療センターの編
上雅史肝炎・免疫研究
センター長らが突き止
めた。遺伝子を探られ
ば、より安全で効果よ

C型肝炎治療 遺伝子が左右

効かない人は型に違い

い治療を進められると、二百万人とされる。治
いう。一四日付の米科 疫は、インターフェロ
学誌「ネイチャー・シエ ン」と抗ウイルス剤の投
ネティクス(種子版) 与が主だが、副作用のう
に発表した。
C型肝炎はウイルス 人に届く多量タイプ
(HCV)感染で起る 場合、効果が出ない人
る。国内の感染者は約 数が少ない。

名市大准教授ら解明

そこでこのタイプの C型肝炎の治療には
患者約三百五十人の血 発熱やうつ病などの副
被を調べたところ、薬 作用の差が伴う。田
が効く人と効かない人 中准教授は「効く人に
とでDNA型に微妙な 差は、遺伝的に治療し、効
違いがあることが分か かない人には不十分な
った。型の違いを調べ 昔病や出芽を繰り返す
れば、約80%の確率で 発症の新薬を待っても
治療効果が出るかと いう人、なぜか治療が
かかるといふ」
と研究「」と題し、

C型肝炎治療の効果 遺伝子型で差

国内研究班

特定の遺伝子の型を「型」と呼ぶ。C型肝炎ウイルスの遺伝子の型は、主に「1型」と「2型」に分かれる。このうち、1型は日本に多く見られる。研究班は、C型肝炎ウイルスの型が治療効果に与える影響を調べるため、国内研究班が実施した。研究班は、C型肝炎ウイルスの型が治療効果に与える影響を調べるため、国内研究班が実施した。

い人が約のいる。約三百五十人の患者の血液を調べたところ、異なる人と効かない人の間で、RNAに似た型が多い(遺伝子多型)が見られた。この多型を調べると、約8割の患者で治療効果が出たことが分かる。研究班は、この結果を踏まえ、治療法を改善する必要があると指摘している。

研究班は、C型肝炎ウイルスの型が治療効果に与える影響を調べるため、国内研究班が実施した。研究班は、C型肝炎ウイルスの型が治療効果に与える影響を調べるため、国内研究班が実施した。

