

200728026A

厚生労働科学研究費補助金
肝炎等克服緊急対策研究事業

テーラーメイド治療を目指した肝炎ウイルス
データベース構築に関する研究
(H19-肝炎-一般-013)

平成 19 年度 総括・分担研究報告書

主任研究者 溝上 雅史

平成 20 (2008) 年 3 月

目 次

I. 総括研究報告書	頁
テラーメイド治療を目指した肝炎ウイルスデータベース構築に関する研究 ……………	1
(名古屋市立大学 溝上 雅史)	
II. 分担研究、研究協力報告書	
1. 肝炎ウイルス統合データベースの構築 ……………	9
(国立遺伝学研究所 五條掘 孝)	
2. ゲノムワイド関連分析による肝炎の宿主遺伝要因の探索……………	17
(東京大学 徳永 勝士)	
3. 肝外病変とくに扁平苔癬と HCV 感染……………	21
(久留米大学 佐田 通夫)	
4. 肝癌浸潤リンパ球、肝癌患者末梢血液単核球細胞の 遺伝子発現プロファイルの特徴……………	23
(金沢大学 本多 政夫)	
5. C型慢性肝炎に対するペグインターフェロン・リバビリン併用療法 における治療効果規定因子のデータマイニング解析……………	25
(武蔵野赤十字病院 黒崎 雅之)	
6. B型肝炎ウイルスの X、コアプロモーター遺伝子変異と肝発癌機序の解明……………	28
(名古屋市立大学 田中 靖人)	
III. 研究成果の刊行一覧……………	33
IV. 研究成果の刊行物・別冊……………	39

I. 総括研究報告書

厚生労働科学研究費補助金（肝炎等克服緊急対策研究事業）
総括研究報告書（平成 19 年度）

テーラーメイド治療を目指した肝炎ウイルスデータベース構築に関する研究

主任研究者：溝上 雅史 名古屋市立大学大学院医学研究科 臨床分子情報医学 教授

研究要旨：ウイルス性肝疾患はホストのヒト側要因と病原体のウイルス側要因の両方の相互関係で病態は決定されるが、ヒト側要因は多くの要因が複雑に絡み合っていることから従来網羅的に研究することが難しかった。肝組織や末梢血液細胞におけるDNAマイクロアレイを用いた網羅的遺伝子発現解析に加えて、今回新しく開発された約90万個のSingle nucleotide polymorphism (SNP) が本邦でも十分使用可能であることを本年度は明らかにした。この結果に基づいて、来年度以降ヒトゲノム倫理委員会の承認を得た全国15施設（2施設は再申請中）の研究協力者の臨床検体を使用して肝炎ウイルス感染に対する発症感受性や薬剤応答性等の個人差と病態との関係を明らかにしていく。一方、今後明らかになるであろう各種情報を統合したデータベースの基盤設計も終了し、来年度以降、本研究班の本来の目的であるテーラーメイド治療への応用の研究が開始可能となった。

A. 研究目的

本邦における慢性肝疾患の原因ウイルスである B 型肝炎ウイルス (HBV) 感染者は約 150 万人、C 型肝炎ウイルス (HCV) 感染者は約 200 万人と推定され、その一部は慢性肝炎から肝硬変・肝細胞癌へ移行するがその機序は不明である。また、HCV にはペグインターフェロン・リバビリン併用療法が導入され、一定の効果は認められているが完治率は 50%以下でその副作用の発現率は高率である。本研究では、肝炎ウイルス感染に対する応答性の個人差に関わるヒト SNP を明らかにし、さらにウイルス側の遺伝子要因も明らかにし、両者を網羅的に収集し統合的にデータベース化することにある。そして最終的に肝炎テーラーメイド治療の確立を目指す。

B. 研究方法

主任研究者である溝上雅史は、本研究の主研究

施設である名古屋市立大学大学院医学研究科ヒトゲノム倫理委員会では本研究の承認の獲得と各施設での本研究の承認の獲得を受け患者同意に基づいて検体採取を開始した。また、田中靖人分担研究者と共同で、各種患者情報をマッチさせた Case-control study から HBV で発がんとの関係が強く示唆されている X~core promoter~precore 領域の全塩基配列を決定することで HBV 発がんに関与する HBV 遺伝子変異の同定を行った。

五條堀孝分担研究者は、肝炎ウイルス統合データベース (HVDB) に本年度新しく発表された HCV と HBV 配列を HVDB ウイルスに加え更新し、ヒト側要因である患者 SNP 情報のスキーマの構築と患者情報匿名化の設計を試みた。

徳永勝士分担研究者は 90 万種の SNP タイピングが測定可能な Affymetrix 社の SNP Array 6.0 が日本人にどれくらい応用可能かどうかの検討と健常者におけるその頻度の測定を行っ

た。

佐田通夫分担研究者は肝炎ウイルスによる肝外病変における粘膜病変の検討を行うために参加頂き、本年度は倫理委員会の審査承認と患者同意および付帯情報の収集にあたった。

本多政夫分担研究者は自施設で開発した DNA マイクロアレイを用いて、HCV に感染している肝硬変患者と肝細胞がん患者の免疫担当細胞と肝細胞がん患者の手術組織でがん部と非がん部の免疫担当細胞における網羅的遺伝子発現パターンを検討した。

黒崎雅之研究協力者はデータマイニング手法を用いて一般臨床検査情報とペグインターフェロン・リバビリン併用療法の効果を検討した。

C. 研究結果

主任研究者である溝上雅史は、本研究に関して名古屋市立大学大学院医学研究科ヒトゲノム倫理委員会の承認を得て、患者の同意に基づいて検体採取を開始した。

そして、田中靖人分担研究者と共同で、HBV 関連肝細胞癌患者 80 人、非肝癌患者 80 人の HBVDNA の X-core promoter-precore 領域の塩基配列を決定し、肝癌に寄与する HBV 変異の同定を試みた。本邦における HBV キャリアは母親からの垂直感染により起こり約 80-90%は成人期に HBe 抗原から HBe 抗体へのセロコンバージョンをおこし、予後良好の無症候性キャリアへと移行する。一方、約 10-20%の症例においては肝機能異常が持続し、慢性肝炎、肝硬変、肝細胞癌へと移行する。これらの臨床経過の違いは生体因子と HBV 因子の 2 つに大別される。生体因子としては各種要因があるが本研究ではヒト SNP を、また HBV 因子としては HBV の遺伝子

変異を原因と考え両者の関連性を明らかにすることにある。そこで、本年度は、HBV genotype C (HBV/C) 症例において、年齢、性別、HBe 抗原陽性率をマッチさせた肝がん患者 (HCC グループ) と対照として非肝がん患者 (non-HCC グループ) の HBV の X-core promoter-precore 領域の全塩基配列を決定し、どの部位の変異が、肝がん進展に関与しているかを検討したところ、basal core promoter (BCP) の変異 (T1762/A1764) に alpha box 1653 番目の C から T への変異 (T1653) や 1753 番目の変異が加わることにより肝がん発生に関与することを明らかにした。来年度以降これらの患者の SNP を測定し、両者を合わせて検討することで、生体因子と HBV 因子の両方の関係が明らかになると思われる。

五條堀孝分担研究者は、肝炎ウイルス統合データベース (HVDB) 構築の一環として、本年度は HCV 配列 59,906 件、HBV 配列 14,375 件を HVDB ウイルスに新規に加え更新した。その結果、本 HVDB に収録できた配列数は HCV61,316 件、HBV16,605 件に増加した。さらに、現在解析中のヒト側要因である患者 SNP 情報を収集、管理、解析するデータベースのスキーマの構築と患者情報匿名化に容易に対応できる患者臨床情報データベースの設計を行った。

徳永勝士分担研究者はまず Affymetrix 社の SNP Array 6.0 を用いて、健常者 200 人のゲノム DNA について 90 万種の SNP タイピングを完了した。さらに、慢性ウイルス性肝疾患患者 144 検体の SNP タイピングが完了した。健常対照群 200 検体の SNP タイピング結果から日本人試料では約 60 万種類の SNP について解析可能であることが明らかとなった。さらに、慢性ウイル

ス性肝疾患患者 96 検体と健常対照群 198 検体を対象としたゲノムワイド関連解析を実施したところ、ヒト 6 番染色体上の MHC 領域 (6q21.31) には統計的に有意な SNP が存在することが明らかとなった。慢性ウイルス性肝疾患患者 200 検体の SNP タイピングが完了し次第、患者群を病態および薬剤応答性に応じて分類したサブグループに分類しゲノムワイド関連解析を行うことにより、宿主側の遺伝要因の候補が検出される可能性が明らかとなった。

佐田通夫分担研究者は肝炎ウイルス感染における肝外病変の国内外の第一人者である。HBV や HCV は肝外病変として扁平苔癬やシェーグレン症候群などの粘膜病変を 10-20%もの患者に引き起こすことが、その発生メカニズム及びどのような人に発生するのかは知られていない。また、HBV や HCV の治療薬であるインターフェロン投与中には高率にこれらの病変は悪化し、またインターフェロン投与中に新たに発生し、しばしばインターフェロン治療の中断を余儀なくされる。このため本研究班では佐田通夫分担研究者に参加頂いた。本年度は本研究の開始にあたり久留米大学倫理委員会の審査で承認を受けた後(承認日 2007 年 8 月 6 日付)、研究体制を確立し 2007 年 10 月 30 日より検体及び付帯情報の収集を開始された。現在(平成 20 年 3 月 30 日)、約 30 例が解析を終了した。そして、来年度中には約 100 例を目標に集積、解析し、扁平苔癬やシェーグレン症候群などの粘膜病変に関与するヒト SNP の同定を行う予定となっている。

本多政夫分担研究者は網羅的遺伝子発現解析の専門家である。肝炎ウイルス感染の最終結末は肝細胞がんであり、その原因を解明し肝細胞

がんを予防することが本研究班の最終目的の一つでもある。その肝細胞がんへの進展は肝炎ウイルスと生体の特に免疫細胞が複雑に絡み、ある患者は肝がんへ、ある患者は進展しないと考えられている。しかしながら、あまりにも多くの因子が絡むためにその要因を特定することは困難であった。近年本多政夫分担研究者らにより DNA マイクロアレイを用いて肝組織において包括的に各種遺伝子発現を極めて精密に検討することが可能となった。そこで、DNA マイクロアレイを用いて肝がんに関与するヒト因子を検討してもらうために本多政夫分担研究者に参加して頂いた。本年度は、HCV に感染している肝硬変患者と肝細胞がん患者の免疫担当細胞と肝細胞がん患者の手術組織でがん部と非がん部の免疫担当細胞を超選択的に取り出し DNA マイクロアレイを用いてそれぞれ包括的に遺伝子発現を解析した。その結果、抗原提示、低酸素および酸化ストレスへの応答、ユビキチン・プロテアゾーム蛋白分解、細胞周期、mRNA プロセスに関連する因子が亢進していたが、興味深いことに肝がん患者 PBMC において発現が亢進した遺伝子の生物学的プロセスと同一であった。このことは、肝がん患者において末梢血液細胞の遺伝子発現解析によって宿主のがんに対する応答を解明できる可能性が示された。さらに、来年度以降のヒト SNP の解析と同時に一緒に解析することで肝がん進展への要因がさらに精密に特定できる可能性が考えられた。

黒崎雅之研究協力者はデータマイニング手法の専門家である。現時点における HCV 感染の根治治療の根幹をなすのはペグインターフェロン・リバビリン併用療法である。しかしながらその効果

は約 50%で高価な薬であるにもかかわらず十分とはいえない。その理由として従来この治療の成否を規定する因子としてウイルス側因子(ウイルス因子)のみが精査されてきたためである。そこで本研究班では、ウイルス(HCV)因子のみならず感染母体である生体因子も考慮して検討することとしたが、その生体因子としては一般臨床検査で検討できる因子(生体因子)とホスト遺伝子で規定される因子(ヒト SNP)に大別される。そこで、本研究班では、この生体因子で検討可能なデータマイニングの第一人者である武蔵野赤十字病院の黒崎雅之研究協力者に参加頂き、今年度はペグインターフェロン・リバビリン併用療法の効果にウイルス因子と生体因子がどのようにかわるかを検討して頂いた。その結果、本年度は、steatosis、LDL-cholesterol、年齢、gamma-GTP、および血糖値などの各種の生体因子の方がウイルス因子よりも、ペグインターフェロン・リバビリン併用療法の治療効果を規定する治療抵抗性因子として強いことを明らかにした。さらに、先の生体因子の影響を最小限にするためにこれらの生体因子を均一化した集団を作成し、それらの中でウイルス因子を検討したところ、HCV NS5A と Core 領域の変異が治療効果を大きく規定することを明らかにした。この事実はすぐにも臨床に応用可能であるが、来年度以降は現在進行中のヒト SNP の解析結果も順次このデータマイニングに投入し、生体因子、ウイルス因子、ヒト SNP 情報の統合的解析を行う予定となっている。

D. 考察

ウイルス性肝疾患は肝炎ウイルス感染症であるのでホストのヒト側要因と病原体であるウイルス側要因の両方の相互関係で病態は決

定される。したがって、これら両方の要因を同時に検討しないことにはウイルス性肝疾患の本体を明らかにできない。しかしながら現在まで主にウイルス側要因に関する研究しかなされてこなかった。その理由として、ヒト側要因は多くの要因が複雑に絡み合っていることが予想されるがそれらを網羅的に解析する手法が存在しなかったからである。

近年、そのヒト側要因を一度に網羅的に約 90 万個の Single nucleotide polymorphism (SNP) で測定することが可能となった。そこで、この SNP 測定を本邦の第一人者である徳永勝士分担研究者にお願いして行って頂いた。まず、Affymetrix 社の SNP Array 6.0 を用いて日本人健常者 200 人のゲノム DNA の検討を行って頂き、日本人試料でも約 60 万種類の SNP について解析可能であることが明らかとなった。

一方、本多政夫分担研究者は自ら開発した DNA マイクロアレイを用いて肝組織において網羅的に解析し肝がん患者は抗原提示、酸化ストレスへの応答などの因子が亢進していることを明らかにし、さらに肝がん患者では末梢血液細胞の遺伝子発現解析によって宿主のがんに対する応答を解明できる可能性を示した。

これらの結果、来年度以降ヒトゲノム倫理委員会の承認を得た全国 15 施設(2 施設は再申請中)の研究協力者の臨床検体を使用して順次、肝炎ウイルス感染に対する応答性(発症感受性および重症化)や薬剤応答性の個人差各種などの病態との関係を明らかにしていくことが可能となった。

一方、PCR 法を始めとする近年の分子生物学の発展は目覚しく、特に各種生物の塩基配列決定の自動化、高速化、簡易化がなされた。その

結果、世界各地より多数の塩基配列が国際版データベース DDBJ/EMBL/GenBank に登録されている。しかしながら、あまりにも多くの塩基配列が報告された結果、専門家以外にはその操作は複雑且つ煩雑になり、手に負えなくなっていた。そこで、我々は肝炎ウイルス専用のデータベースを 1996 年から作成し現在まで維持してきた。また、インターネットに見られる電子情報処理技術も急速に進歩した。そこで、この両者の長所を合わせヒト SNP とウイルスゲノム解析に基づいた統合型肝炎ウイルスデータベースを構築し、肝炎テーラーメイド治療の確立を目指すこととし、この分野の第一人者である五條堀孝分担研究者にお願いし、ウイルス配列データベースの更新と患者 SNP データベース及び患者情報データベースの設計を行って頂いた。

以上の結果より、今年度で本研究班の基盤は確立され、来年度以降の実際の臨床検体を使用した実施が可能となった。

E. 結論

本年度内に、本研究班の一つの根本であるヒト側要因を網羅的に解析する手法が本邦でも使用可能であり、またその検体の提供を受ける全国 13 施設 (2 施設は再申請中) でのヒトゲノム倫理委員会の審査も終了した。さらに、ウイルス側要因の臨床的、ウイルス学的検討も順調に進んでいる。そして、これらを統合するデータベースの基盤設計も終了し、来年度以降本研究班の本来の目的であるテーラーメイド医療への応用が進められることが明らかとなった。

F. 研究発表

1. 論文発表

1. Lok AS, Zoulim F, Locarnini S, Bartholomeusz A, Ghany MG, Pawlotsky JM,

- Liaw YF, Mizokami M, Kuiken C; Hepatitis B Virus Drug Resistance Working Group. Antiviral drug-resistant HBV: standardization of nomenclature and assays and recommendations for management. *Hepatology*. 2007 Jul;46 (1):254-65. Review.
2. Shin-i T, Tanaka Y, Tateno Y, Mizokami M. Development and Public Release of Comprehensive Hepatitis Virus Database. *Hepatol Res*. 2007 in press.
3. Sakamoto T, Tanaka Y, Simonetti J, Osiowy C, B_rresen ML, Koch A, Kurbanov F, Sugiyama M, Minuk GY, McMahon BJ, Joh T, Mizokami M. Classification of hepatitis B virus genotype B into two major types based on characterization of a novel subgenotype in the Arctic indigenous populations. *J Infect Dis*. 196 (10):1487-92. 2007.
4. Yano Y, Truong BX, Seo Y, Kato H, Miki A, Tanaka Y, Mizokami M, Kagawa A, Miyazaki H, Azuma T, Kasuga M, Hayashi Y. A Japanese case of hepatitis B virus genotypes C/D hybrid. *Hepatol Res*. 2007 in press.
5. Kusakabe A, Tanaka Y, Orito E, Sugauchi F, Kurbanov F, Sakamoto T, Shinkai N, Hirashima N, Hasegawa I, Ohno T, Ueda R, Mizokami M. A weak association between occult HBV infection and non-B non-C hepatocellular carcinoma in Japan. *J Gastroenterol*. 42:298-305. 2007.
6. Nukaya H, Ohno T, Sakakibara K, Kato A, Hasegawa I, Matunaga S, Endo M, Tanaka Y, Hirashima N, Tanaka Y, Orito E, Joh T, Mizokami M. Accidental exposure to HCV

antibody-positive blood in hospital and pre-emptive one-shot interferon alpha-2b treatment. *Hepatol Res.* 37 (3) :179-85, 2007.

7. Fung J, Lai CL, Yuen JCH, Wong DKH, Tanaka Y, Mizokami M, Yuen MF. Adefovir dipivoxil monotherapy and combination therapy with lamivudine for the treatment of chronic hepatitis B in an Asian population. *Antiviral Therapy* 12: 41-46, 2007.

2. 学会発表

1. Sugauchi F, Orito E, Y. Tanaka, Ozasa A, Kang J, Toyoda J, Yotsuyanagi H, Iino S, Kuramitsu T, Suzuki K, Tanaka E, Akahane Y, Izumi N, Inoue K, Kakumu S, Tomita E, Murawaki Y, Hino K, Onji M, Yatsushashi H, Sata M, Okanoue T, Nishiguchi S, Miyakawa Y, Mizokami M. Influence of hepatitis B virus genotypes and G1896A mutation on fulminant outcome of acute infection. 17th Asian Pacific Association for the Study of the Liver Disease (APASL). March 27-30 2007. Kyoto.

2. Shinkai N, Tanaka Y, Ito K, Mukaide M, Hasegawa I, Asahina Y, Izumi N, Yatsushashi H, Orito E, Joh T, Mizokami M. Influence of hepatitis B virus X, core promote mutations on hepatocellular carcinoma among patients with subgenotype C2. 17th APASL Conference. March 27-30 2007. Kyoto.

3. Mukaide M, Tanaka Y, Orito E, Yuen MF, Ito K, Kurbanov F, Sugauchi F, Asahina Y, Izumi N, Kato M, Lai CL, Ueda R, Mizokami M. Specific Mutations in Enhancer II/Core

Promoter of Hepatitis B Virus Subgenotypes C1/C2 Increase the Risk of Hepatocellular Carcinoma. 17th APASL Conference. March 27-30 2007. Kyoto.

4. Shinkai N, Tanaka Y, Ito K, Mukaide M, Hasegawa I, Asahina Y, Izumi N, Yatsushashi H, Orito E, Joh T, Mizokami M. Influence of hepatitis B virus X, core promote mutations on hepatocellular carcinoma among patients with subgenotype C2. 14th Japan-Korea International Hepatitis Meeting. June16-17 2007. Fukuoka.

5. Kurbanov F, Tanaka Y, Elkady A, Khan A, Oyunsuren Ts, Sanduijav R, Khajidsuren O, Dagvadorj B, Mizokami M. Viral factors associated with Hepatocellular carcinoma (HCC) in Mongolian chronic hepatitis patients. 14th Japan - Korea International Hepatitis Meeting. June16-17 2007. Fukuoka.

6. Masaya Sugiyama, Yasuhito Tanaka, Satoru Takahashi, Tomoyuki Shirai, Isao Maruyama, Takashi Shimada and Masashi Mizokami. Virological and histopathological characteristics among hepatitis B virus genotypes in uPA/SCID mice with human hepatocytes. 2007 International meeting the molecular biology of Hepatitis B virus. 2007. Roma.

7. Tadasu Shin-i, Masashi Mizokami. Update on HCV Sequence Databases - Japanese HVDB (in HCV classification and nomenclature group 2nd meeting). 14th International Symposium on Hepatitis C Virus & related Viruses. 2007.

Glasgow, U. K.

8. 田中靖人, Yuen Man-Fung, 溝上雅史. エンテカビル治療効果及び耐性変異に関する検討. 第11回日本肝臓学会大会. 平成19年10月18-19日. 神戸.
9. 日下部篤宣, 田中靖人, 狩野吉康, 古賀郁利子, 坂井田功, 佐久川廣, 鈴木一幸, 住野泰清, 田中榮司, 長田成彦, 山田剛太郎, 持田智, 溝上雅史. B型肝炎劇症化に寄与するウイルス因子の検討. 第11回日本肝臓学会大会. 平成19年10月18-19日. 神戸.
10. Masaya Sugiyama, Yasuhito Tanaka, Isao Maruyama, Takashi Shimada and Masashi Mizokami. Differences of early dynamics and liver damage among hepatitis B virus genotypes in uPA/SCID mice with human hepatocytes. 58th Annual Meeting of the American Association for the Study of Liver Diseases (AASLD). 2007. Boston, MA.
11. Kurbanov F, Tanaka Y, Maruyama I, Shimada T, Mizokami M. Interferon Sensitivity of Natural Recombinant RF1_2k/1b HCV strain in chimeric uPA/SCID mice. 58th AASLD. 2007. Boston, MA.
12. Kurbanov F, Tanaka Y, Elkady A, Oyunsuren Ts, Sanduijav R, Khajidsuren O, Dagvadorj B, Mizokami M. Why is HCC incidence so high in Mongolia Joint Research Workshop: International comparative-epidemic study on the relationship among persistent infection of hepatitis virus and hepatocellular carcinoma between Japan and other countries. 2008. Hiroshima.

G. 知的所得権の所得状況

1. 特許取得
なし
2. 実用新案登録
なし

II. 分担研究、研究協力報告書

厚生労働科学研究費補助金（肝炎等克服緊急対策研究事業）
分担研究報告書（平成 19 年度）

テラーメイド治療を目指した肝炎ウイルスデータベース構築に関する研究

分担研究者：五條掘 孝 国立遺伝学研究所 生命情報・DBJ 研究センター
遺伝情報分析研究室 教授

分担研究課題：肝炎ウイルス統合データベースの構築

研究要旨：宿主（ヒト）側と肝炎ウイルス側双方の要因に注目し、広範なサンプルより両データを収集し、患者より得られた臨床情報を加えて多次元的にデータベース化し解析する。本研究では、ヒト側要因である患者SNPs情報を収集、管理、解析するデータベースが必要となるため、その設計を目的として調査および情報収集を行い、DBスキーマを構築する。併せて、ウイルス側要因のデータベースであるHVDB (Hepatitis Virus DataBase) を継続して更新し、研究者の利用に供する。

A. 研究目的

本研究は、ヒト側・ウイルス側双方の要因を統合的に解析することが特徴であり、それに必要な基盤として、両データを網羅的に収集し相互参照的解析を可能とするデータベースを構築する。

B. 研究方法

ヒト側要因である患者 SNPs データは、Affymetrix GeneChip により検出されるため、その出力データ形式の調査を行う。また、参照情報として使用する dbSNP データベースの内部構造を調査する。これらの結果を踏まえて、DB スキーマの定義を行う。

C. 研究結果

前述の調査を行った結果、SNP 情報と患者臨床情報を分離させる必要性が判明した。そこで、データベース全体をウイルス配列 DB、患者 SNP DB、患者臨床情報 DB の 3 要素で構成し、患者臨床情報 DB が前の 2DB を連絡する形式にスキーマを定義した。（別添書類参照）

併せて、従来より公開している HVDB を更新した結果、HCV 配列 61, 316 件、HBV 配列 16, 605 件に増加した。

D. 考察

本データベースは 3 要素構成で定義されているため、患者臨床情報 DB を重点的に保護することで、患者情報匿名化に容易に対応でき、必要な機密性を実現できると考えられる。

E. 結論

本研究で対象となる SNP データの構造を調査し、それを踏まえて肝炎ウイルス統合データベースの設計を行った。次年度以降、データ収集・解析に実際に供することが可能である。

F. 研究発表

1. 論文発表

1. Murakami, K., Imanishi, T., **Gojobori, T.** and Nakai, K. (2008). Two different classes of co-occurring motif-pairs found by a novel visualization method in

- human promoter regions. *BMC Genomics* (in press)
2. Nakagawa, S., Niiimura, Y., **Gojobori, T.**, Tanaka, H. and Miura, K. (2008). Diversity of preferred nucleotide sequences around the translation initiation codon in eukaryote genomes. *Nucleic Acids Res.* 36(3), 861-871
 3. Genome Information Integration Project and H-Invitational 2 Consortium: Yamasaki, C., Imanishi, T., **Gojobori, T.** et al. (2008). The H-Invitational Database (H-InvDB), a comprehensive annotation resource for human genes and transcripts. *Nucleic Acids Res.* 36(Database issue), D793-799.
 4. Matsuya, A., Sakate, R., Kawahara, Y., Koyanagi, KO., Sato, Y., Fujii, Y., Yamasaki, C., Habara, T., Nakaoka, H., Todokoro, F., Yamaguchi, K., Endo, T., Oota, S., Makalowski, W., Ikeo, K., Suzuki, Y., Hanada, K., Hashimoto, K., Hirai, M., Iwama, H., Saitou, N., Hiraki, AT., Jin, H., Kaneko, Y., Kanno, M., Murakami, K., Noda, AO., Saichi, N., Sanbonmatsu, R., Suzuki, M., Takeda, J., Tanaka, M., **Gojobori, T.**, Imanishi, T. and Itoh, T. (2008). Evola: Ortholog database of all human genes in H-InvDB with manual curation of phylogenetic trees *Nucleic Acids Res.* 36(Database issue), D787-792.
 5. Tateno, Y., Sugawara, H., Ogasawara, O., Okubo, K. and **Gojobori, T.** (2008). DDBJ with New System and Face. *Nucleic Acids Res.* 36(Database issue), 22-24.
 6. Makino, T. and Gojobori, T. (2007). Evolution of Protein-Protein Interaction Network. *Gene and Protein Evolution. Genome Dynamics* Vol. 3, 13-29.
 7. Tanaka, Y., Hanada, K., Hanabusa, H., Kurbanov, F., **Gojobori, T.** and Mizokami, M. (2007). Increasing genetic diversity of hepatitis C virus in hemophiliacs with human immunodeficiency virus coinfection. *J. Gen. Virol.* 88(Pt 9), 2513-2519.
 8. Kubota, R., Hanada, K., Furukawa, Y., Arimura, K., Osame, M., **Gojobori, T.** and Izumo, S. (2007) Genetic Stability of Human T Lymphotropic Virus Type I despite Antiviral Pressures by CTLs.. *Journal of Immunol.* 178(9), 5966-5972.
 9. Liu, QX., Nakashima-Kamimura, N., Ikeo, K., Hirose, S. and **Gojobori, T.** (2007). Compensatory Change of Interacting Amino Acids in the Coevolution of Transcriptional Coactivator MBF1 and TATA-Box Binding Protein TBP. *Mol. Biol. Evol.* 24(7), 1458-1463.
 10. Osato, N., Suzuki, Y., Ikeo, K. and **Gojobori, T.** (2007). Transcriptional interferences in cis natural antisense transcripts of human and mouse. *Genetics* 176, 1299-1306.
 11. Hanada, K., Tanaka, Y., Mizokami, M., **Gojobori, T.** and Alter, HJ. (2007). A reduction in selective immune pressure during the course of chronic hepatitis C correlates with diminished biochemical evidence of hepatic inflammation. *Virology* 361, 27-33.
 12. Takeda, JI., Suzuki, Y., Nakao, M., Kuroda, T., Sugano, S., **Gojobori, T.** and Imanishi, T. (2007). H-DBAS: Alternative Splicing Database of Completely Sequenced and Manually Annotated Full-length cDNAs Based on H-Invitational. *Nucleic Acids Res.* 35 (Database issue), D104-109.
2. 学会発表
1. **T. Gojobori** (2007) “Genetic Polymorphisms as a phase of evolutionary process, and the future of population genomics” (基調講演), 5th International Bio data Interoperability Conference, International Symposium “State-of-the-art of polymorphism study”,

National Institute of Advanced Industrial
Science and Technology (Tokyo) 9月26日

2. 五條堀 孝 (2008)「遺伝子転写制御ネットワークの全容理解に向けての情報基盤の構築-ゲノムネットワークから見えてきたもの-」、内部交流セミナー、遺伝学研究所（三島）1月25日
3. 五條堀 孝 (2008)「ゲノム生物学における中立進化論の役割と課題」、学術講演会「分子進化の中立論40周年」、東京大学本郷キャンパス（東京）2月17日
4. 五條堀 孝 (2008)「病態発現機構の解明に役立つバイオインフォマティクスの研究開発」、東京医科歯科大学難治疾患研究所、病態発現機構客員研究部門、客員部門研究報告セミナー、東京医科歯科大学難治疾患研究所（東京）3月14日

G. 知的所得権の所得状況

1. 特許取得
特になし。
2. 実用新案登録
特になし。
3. その他
特になし。

1. 概要

本ファイルには、肝炎ウイルス統合データベースをRDB(関係データベースシステム)上に構築する際のスキーマ(テーブル構造定義)が記述されています。

2. 各シートの説明

2. 1. 共通事項

unique key そのテーブルにおける一次キー
outer key 他のテーブルのデータと結合する際に使用されるキー

2. 2. 「SNP」シート

ホスト側要因、すなわち、患者SNPに関連するデータを格納するテーブル群が定義されています。
患者SNPの検出結果、頻度解析の結果、GeneChipの構造、各SNPに関連する公知情報へのリンクが含まれます。

↳

2. 3. 「患者」シート

患者に関連する情報を格納するテーブル群が定義されています。
患者個人情報概要、診断履歴、投薬履歴、サンプル採取履歴が含まれます。

2. 4. 「配列」シート

ウイルス側要因、すなわち、ウイルスゲノム配列に関連するデータを格納するテーブル群が定義されています。
ウイルス種類、サブタイプ、遺伝子座、核酸/アミノ酸配列、変異情報が含まれます。

2. 5「相関図」シート

上記で定義されているテーブル群の相互関係が図示されています。

chip_layout	GeneChipのレイアウト			
	id_snp	int	unique key	SNPのID
	id_chip	int	outer key	GeneChipの種類(ID)
	SNP	char(64)		検出されるSNPの記述子、同一チップ内ではユニーク
	position	int		SNPプロローブのチップ上の位置、同一チップ内ではユニーク
	acc_db_snp	char(32)		dbSNPに既登録の場合、そのアクセッション
acc_isnp	char(32)		isNPに既登録の場合、そのアクセッション	

chip_name	GeneChipの種類			
	id_chip	int	unique key	GeneChip種類ID
	name	char(64)		GeneChipの種類を要す記述子、モデル番号など
	num_SNP	int		検出できるSNPの数
	comment	text		付加情報

chip_experiment	SNP検出実験			
	id_expr	int	unique key	実験ID
	id_sample	int	outer key	使用したサンプル(ID)
	date_expr	date		実験日時
	id_chip	int	outer key	使用したGeneChip (ID)
	computed_gender	char(1)		自動判定された性別、M/F
	SNP_call_rate	double		判定できたSNPの率
	SNP_call_AA_rate	double		AA型と判定された率
	SNP_call_AB_rate	double		AB型と判定された率
	SNP_call_BB_rate	double		BB型と判定された率
	num_SNP_QC	int		QCに使用されたSNP数
	QC_intensity	double		intensityの中央値(メジアン)
	QC_MCR	double		MPAMアルゴリズムによる判定率
	QC_MDR	double		MPAMアルゴリズムによる判定率

chip_experiment_snp	SNP検出実験、SNP単位のデータ			
	id_expr_snp	int	unique key	検出SNPのID
	id_expr	int	outer key	対応する実験(ID)
	id_snp	int	outer key	対応するSNP(ID)
	result	char(2)		判定結果、AA/AB/BB
	intensity	double		intensity
	score	double		confidence score

chip_analysis	頻度分布解析			
	id_freq	int	unique key	頻度分布解析のID
	id_exgroup_tgt	int	outer key	対象グループ(ID)
	id_exgroup_ref	int	outer key	対照グループ(ID)
	num_SNP_valid	int		有意な差のあったSNP

chip_analysis_snp	頻度分布解析、SNP単位のデータ			
	id_freq_snp	int	unique key	解析SNPのID
	id_freq	int	outer key	対応する解析結果(ID)
	SNP_AA_tgt	int		対象グループにおけるAA型の数
	SNP_AB_tgt	int		対象グループにおけるAB型の数
	SNP_BB_tgt	int		対象グループにおけるBB型の数
	SNP_AA_ref	int		対照グループにおけるAA型の数
	SNP_AB_ref	int		対照グループにおけるAB型の数
	SNP_BB_ref	int		対照グループにおけるBB型の数
	chisquare	double		検定値

experiment_group	実験結果グループ			
	id_exgroup	int	unique key	グループ番号ID
	id_expr	int	outer key	実験(ID), patient_groupの各患者の最新検体の最新実験
	id_group	int	outer key	グループ定義(ID)

publication	関連する論文			
	id_pub	int	unique key	論文ID
	pubmed	int		PubMed ID
	id_snp	int	outer key	対応するSNP (ID)

テーブル名	説明	フィールド名	型	キー種類	説明
patient	患者情報	id_patient	int	unique key	患者ID
		name	char(64)		患者の識別子
		id_hospital	int	outer key	病院 (ID)
		gender	char(1)		性別、M/F
		birthyear	int		生年
		id_place_birth	int	outer key	居住地(ID)
		id_place_infect	int	outer key	感染地(ID)
		id_route	int	outer key	感染経路(ID)
		infectyear	int		推定感染年
		comment	text		コメント
dict_hospital	病院辞書	id_hospital	int	unique key	病院ID
		name	char(64)		病院名
dict_place	地名辞書	id_place	int	unique key	地名ID
		name	char(64)		地名、国名もしくは県名
dict_route	感染経路辞書	id_route	int	unique key	感染経路ID
		name	char(64)		感染経路
diagnosis	診断情報	id_diag	int	unique key	診断ID
		id_patient	int	outer key	患者(ID)
		date_diag	date		診断日
		status	char(64)		症状
		comment	text		コメント
treatment	治療情報	id_treat	int	unique key	治療ID
		id_patient	int	outer key	患者(ID)
		date_treat_from	date		治療開始日
		date_treat_to	date		治療終了日
		id_drug	int	outer key	投与薬剤(ID)
		dose	int		投与量
		comment	text		コメント
dict_drug	薬剤辞書	id_drug	int	unique key	薬剤ID
		name	char(64)		薬剤名
specimen	検体取得情報	id_sample	int	unique key	サンプルID
		id_patient	int	outer key	患者(ID)
		date_ext	date		サンプル収集日
		volume	int		サンプル量
		viral_load	int		ウイルス量
		comment	text		コメント
patient_group	患者グループ	id_member	int	unique key	グループメンバーID
		id_patient	int	outer key	患者(ID)
		id_group	int	outer key	グループ定義(ID)
dict_patient_group	患者グループ定義の辞書	id_group	int	unique key	グループID
		name	char(64)		名称
		definition	char(255)		グループ定義

フィールド名 説明 フィールド名 型 ユニーク制説明

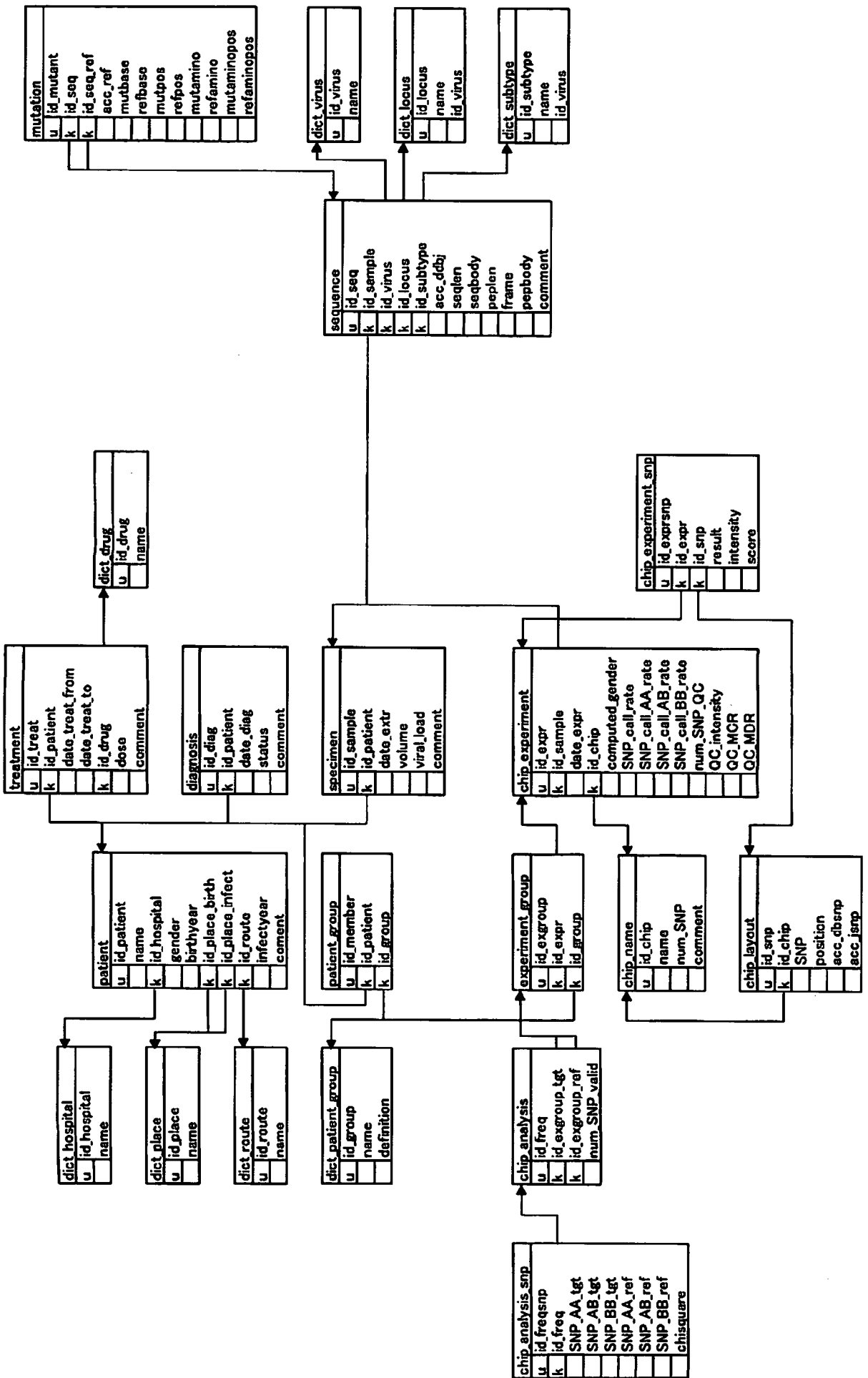
sequence	核酸配列	id_seq	int	unique key	配列ID
		id_sample	int	outer key	由来するサンプルID
		id_virus	int	outer key	ウイルス種類ID
		id_locus	int	outer key	遺伝子座ID
		id_subtype	int	outer key	サブタイプID
		acc_ddbj	char(32)		DBJ既登録データの場、そのアクセッションID
		seqlen	int		配列長
		seqbody	text		配列
		peplen	int		アミノ酸配列長
		frame	int		翻訳フレーム
		pepbody	text		アミノ酸配列
		comment	text		コメント

dict_virus	ウイルス種類の辞書	id_virus	int	unique key	ウイルスID
		name	char(64)		名称

dict_locus	遺伝子座の辞書	id_locus	int	unique key	遺伝子座ID
		name	char(64)		名称
		id_virus	int	outer key	対応するウイルスID

dict_subtype	ウイルスのサブタイプの辞書	id_subtype	int	unique key	サブタイプID
		name	char(64)		名称
		id_virus	int	outer key	対応するウイルスID

mutation	変異	id_mutant	int	unique key	変異ID
		id_seq <td>int</td> <td>outer key</td> <td>変異の見つかった配列ID</td>	int	outer key	変異の見つかった配列ID
		id_seq_ref <td>int</td> <td>outer key</td> <td>参照配列ID</td>	int	outer key	参照配列ID
		acc_ref <td>char(32)</td> <td></td> <td>参照配列アクセッション(公開配列の場合)</td>	char(32)		参照配列アクセッション(公開配列の場合)
		mutbase <td>char(3)</td> <td></td> <td>変異塩基、フレームに合わせた3ベースで記述</td>	char(3)		変異塩基、フレームに合わせた3ベースで記述
		refbase <td>char(3)</td> <td></td> <td>それに対応する参照配列上の3ベース</td>	char(3)		それに対応する参照配列上の3ベース
		mutpos <td>int</td> <td></td> <td>変異位置の先頭</td>	int		変異位置の先頭
		refpos <td>int</td> <td></td> <td>それに対応する参照配列上の位置</td>	int		それに対応する参照配列上の位置
		mutamino <td>char(1)</td> <td></td> <td>変異アミノ酸</td>	char(1)		変異アミノ酸
		refamino <td>char(1)</td> <td></td> <td>それに対応する参照配列上のアミノ酸</td>	char(1)		それに対応する参照配列上のアミノ酸
		mutaminopos <td>int</td> <td></td> <td>変異位置(アミノ酸配列上)</td>	int		変異位置(アミノ酸配列上)
		refaminopos <td>int</td> <td></td> <td>それに対応する参照配列上の位置</td>	int		それに対応する参照配列上の位置



厚生労働科学研究費補助金（肝炎等克服緊急対策研究事業）
分担研究報告書（平成 19 年度）

テーラーメイド治療を目指した肝炎ウイルスデータベース構築に関する研究

分担研究者：徳永 勝士 東京大学大学院医学系研究科人類遺伝学 教授

分担研究課題：ゲノムワイド関連分析による肝炎の宿主遺伝要因の探索

研究要旨：肝炎ウイルスに感染した宿主側の病態促進因子及び薬剤応答性を規定する遺伝要因の解明を目的として、90万種類以上のSNPを対象としたゲノムワイド関連分析を行う。現在までに、健常対照群200検体および慢性ウイルス性肝疾患患者144検体のSNPタイピングが完了し、平均コール率はそれぞれ99.71%、99.58%となった。健常対照群200検体のSNPタイピング結果から日本人試料では約60万種類のSNPが統計解析可能であることが明らかとなり、慢性ウイルス性肝疾患患者96検体と健常対照群198検体を対象としたゲノムワイド関連解析を実施したところ、ヒト6番染色体上のMHC領域（6q21.31）には p 値 < 0.05 となるSNPが期待値の倍以上存在することが明らかとなった。慢性ウイルス性肝疾患患者200検体のSNPタイピングが完了し次第、患者群を病態および薬剤応答性に応じて分類したサブグループごとにゲノムワイド関連分析を行うことにより、宿主側の遺伝要因の候補が多数検出されることが期待される。

A. 研究目的

B型肝炎ウイルスおよびC型肝炎ウイルスに感染した宿主を対象としたゲノムワイド関連分析を行うことにより、宿主側の肝病態進展に寄与する遺伝因子、治療効果に寄与する遺伝因子、ウイルス感染感受性に寄与する遺伝因子を探索することを目的とする。

B. 研究方法

90万種類以上のSNP解析用プローブが搭載されたAffymetrix Genome-Wide Human SNP Array 6.0（以下、Affymetrix SNP6.0）を用いて、慢性ウイルス性肝疾患患者200検体および健常対照群200検体のSNPタイピングを行い、患者群を病態および薬剤応答性に応じてサブグループに分類してゲノムワイド関連分析を行う。

Affymetrix SNP6.0は、制限酵素によるゲノムDNAの断片化とマイクロアレイによるタイピング

の手法に改良を加えることにより、大規模なタイピングを行える手法として確立された。解析対象となるSNPは、公共のSNPデータベースおよびPerlegen社に登録されている約220万種のSNPから遺伝学的情報量が最大化されるように、また連鎖不平衡やHapMapプロジェクトからの情報も考慮して選択された約44万種のSNPに、Tag SNPs、X染色体およびY染色体に存在するSNPsなどを加えた90万種類以上のSNPである。

Affymetrix SNP6.0によるSNPタイピングは、ゲノムの複雑さを低減しマイクロアレイへのハイブリダイゼーション効率を上げるための酵素反応ステップと、洗浄・染色装置およびマイクロアレイ用スキャナーを用いた検出ステップに分けることができる（図1）。90万種のSNPタイピングは、2種類の制限酵素（Sty I、Nsp I）を用いて実現される。制限酵素によるゲノムDNAの断