厚生労働科学研究費補助金（新型インフルエンザ等新興・再興感染症研究事業）
分担研究年度終了報告書

迅速・網羅的病原体ゲノム解析法を基盤とした感染症対策ネットワーク構築に関する研究
～病原体網羅遺伝子解析を基盤にしたプロテオーム解析による抗原解析と新規病原体検査法の開発

研究分担者　梁 明秀　横浜市立大学医学部微生物学
研究協力者　松永智子　横浜市立大学医学部微生物学

研究要旨　近年、次世代核酸シーケンサを用いた新興・再興感染症に対するメタゲノム解析が進歩し、新たな病原体の同定や病因の解明が進んでいる。本研究プロジェクトでは、網羅遺伝子解析を基盤とした病原体のプロテオーム解析を実施し、それを基盤とした病原体に対する新規検査診断法の開発を目指している。網羅的遺伝子情報に基づくウイルス抗原の解析やウイルスタンパク質の機能解析等について、コムギ無細胞タンパク質合成系を用いたプロテオミクスを活用する。本年度は新規病原体 Trichodysplasia spinulosa-associated polyomavirus (TSV) の VP 1 タンパク質を、コムギ無細胞系を用いて作製し、本抗原に対するモノクローナル抗体の作製を試みた。本抗体はウエスタンブロットにおいて VP1 タンパク質およびウイルス様粒子を特異的に認識した。本抗体を用いることで TSV に対する新たな検査法や診断法の開発に結びつくものと考えられる。

A．研究目的

感染症の疑いのある不明疾患やバイオテロ、新興・再興感染症などによるアウトブレイク対策のための迅速・網羅的病原体解析法を基盤とした感染症対策ネットワークシステムの構築が重要である。一方で、感染症の疑いのある不明疾患等のための迅速・網羅的病原体解析法として、ウイルス抗原の検出や血清中の抗ウイルス抗体の測定法を整備することが必要となる。

次世代シークエンシングの進歩に伴ってメタゲノム研究の分野は大きく発展し、不明感染症の病原体由来のゲノム断片を多数検出することが可能である。しかしながら、疾患検体に存在するゲノム断片のみでは、当該病原体の疾患病因との関与について確定することは難しく、核酸検査と平行して疾患臓器における病原体抗原の存在および宿主血清中の病原体特異的抗体の存在を証明すべきである。そのためには、病原体網羅遺伝子解析を基盤にしたプロテオーム解析による抗原解析と新規病原体検査法の開発が必須である。

近年、新興・再興感染症に対する抗原・抗体診断法の開発が進んでいる。しかしながら、従来の手法は、大腸菌や培養細胞への遺伝子導入によりウイルスタンパク質の合成が基盤であり、細胞毒性が強く、かつ可溶性の低いウイルス抗原タンパク質の作製には不向きであった。今回、我々は、コムギ無細胞タンパク質合成系を用いて新規病原体の可溶化全長タンパク質を作製し、これを抗原として用いることで、免疫学的診断に利用可能な高品質のモノクローナル抗体の作製を行った。

B．研究方法

1. コムギ無細胞系による TSV ウイルスタンパク質の合成

Trichodysplasia spinulosa-associated polyomavirus (TSV)がコードする VP1 遺伝子を PCR 法を用いて増幅し、無細胞タンパク質発現ベクター pEU-bls-S1 (bls; biotin ligation site GLNDIFEAQKIEWHE, S1: linker sequence LHPPPPRIS) に導入した。作製した pEU ベクターを鋳型に SPu primer 及び AODA2303 primer を用いて PCR 法により転写鋳型を作製し、SP6 polymerase を用いた転写反応により mRNA を合成、続いてコムギ無細胞合成系・重層法によりタンパ

ク質合成を行った。タンパク質のビオチン化は、下層にコムギ無細胞系で合成した biotin ligase 1μl（～50ng/μl）および 終濃度 0.5μM Biotin を加えることにより行った。免疫用抗原の作製においては、His タグを付加した VP1 タンパク質を、ウイルスタンパク質の可溶化を亢進させるために界面活性剤である Brij35（0.5%）存在下にて合成後、Ni-sepharose ビーズにタンパク質を吸着させた。カラムを 8 M の尿素を含む洗浄液にて 2 回洗浄した後、500mM イミダゾールバッファーを用いて精製タンパク質を抽出した。

2. モノクローナル抗体の作製

コムギ無細胞系を用いて合成したウイルス抗原タンパク質に keyhole limpet hemocyanin(KLH)を共有結合させたものを免疫源とした。これを 6 週齢の Balb/c マウスに foot pad法で2週間置きに1ケ月間免疫した。免疫後、マウス脾細胞を採取し、マウスミエローマ細胞である SP2/O と PEG 法で融合しハイブリドーマを作製した。

3. ELISA 法

96 well plate に終濃度で 50 ng/well の抗原を一晩コートした。反応後、バッファー液を除きブロッキング剤を加えて室温で 1 時間静置した。その後 PBS で 3 回洗浄後、各倍希釈したハイブリドーマ上清を加え室温で1時間反応させた。次に PBS で 3 回洗浄後 HRP 標識抗マウス IgG 抗体を室温で 1 時間反応させた。これを PBS で 3 回洗浄後発色基質を加え、30 分後 1 N 硫酸で反応を停止させマイクロプレートリーダーで吸光度を測定した。

（倫理面への配慮）

本研究において、遺伝子組換え実験を用いることから、研究者が所属する機関の組換え DNA 実験安全委員会、バイオセーフティ委員会、動物実験委員会、医学研究倫理委員会等の承認・認可を得て実験を行った。また、本年度はヒト検体を使用した実験を実施していないが、臨床サンプルの解析及びデータの公表にあたっては、倫理委員会の規則にのっとり、当該患者（感染者）の同意を得る予定である。

C．研究結果

1. TSV-VP1 タンパク質の合成

コムギ無細胞タンパク質合成系を用いて作製した TSV-VP1 タンパク質は、概ね不溶化タンパク質であった。しかしながら、翻訳反応液中に界面活性剤の Brij35 またはリポソームを添加することで、可溶化率が顕著に亢進した。これらの可溶化タンパク質を精製し、モノクローナル抗体作製のための抗原とした。

2. TSV-VP1 モノクローナル抗体の作製

上記にて作製した抗原タンパク質を Balb/c マウスに foot pad 法で 2 週間置きに 2 回免疫した。1 ケ月後マウスの脾細胞を採取し、マウスミエローマ細胞である SP2/0 と PEG 法で融合し 48 種類のハイブリドーマを得た。1 次スクリーニングとして ELISA 法を用いて活性を確認した結果、48 種類のうち、18 種類のハイブリドーマ上清が活性保持していることが明らかとなった。

3. ウェスタンブロット解析による抗原認識の確認

次に上記にて選択した 18 種類のハイブリドーマ上清がウェスタンブロット解析に使用できるか否かについて検討した。VP1 タンパク質が、培養細胞において発現がほとんど見られなかったため、リコンビナント VP1 タンパク質を抗原として用いた。その結果 18 種類すべてのハイブリドーマ上清がウエスタンブロット解析において TSV-VP1 タンパク質を認識できることが確認された。また、VP1 を有するウイルス様粒子を用いた解析においても同様に VP1 タンパク質を検出することができた。現在本抗体の免疫組織化学染色に応用可能かどうかについて検討中である。

D．考察

本研究において、合成した可溶性 TSV-VP1 タンパク質は、機能および構造が保持されている状態であると考えられることから、生体内で誘導される構造を認識する抗体の検出も可能であると示唆される。また、全長タンパク質を用いることで、様々なエピトープに対する抗体が検出されることが期待できる。異なるエピトープを認識する抗体を取得することで、ELISA やイムノクロマト法などの免疫学的アッセイ法の開発に結びつくものと考えられる。

E. 結論

コムギ無細胞タンパク質合成システムを活用して作製した全長 TSV-VP1 タンパク質の精製に成功した。また、本抗原に対するモノクローナル抗体を作製した。今後は、免疫組織化学や免疫沈降法などに有用な抗体クローンを選別することで、TSV の診断や病態解析に貢献できる抗体が得られるものと考えられる。

G. 研究発表

1. 論文発表

(1) Kimura H, Yoshizumi M, Ishii H, Oishi K, Ryo A. Cytokine production and signaling pathways in respiratory virus infection. *Front Microbiol*. 2013 Sep 17;4:276.

(2) Kushibuchi I, Kobayashi M, Kusaka T, Tsukagoshi H, Ryo A, Yoshida A, Ishii H, Saraya T, Kurai D, Yamamoto N, Kanou K, Saitoh M, Noda M, Kuroda M, Morita Y, Kozawa K, Oishi K, Tashiro M, Kimura H. Molecular evolution of attachment glycoprotein (G) gene in human respiratory syncytial virus detected in Japan 2008-2011. *Infect Genet Evol*. 2013 Aug;18:168-73.

(3) Ishigami T, Abe K, Aoki I, Minegishi S, Ryo A, Matsunaga S, Matsuoka K, Takeda H, Sawasaki T, Umemura S, Endo Y. Anti-interleukin-5 and multiple autoantibodies are associated with human atherosclerotic diseases and serum interleukin-5 levels. *FASEB J*. 2013 Sep;27(9):3437-45.

(4) Kiyota N, Kushibuchi I, Kobayashi M, Tsukagoshi H, Ryo A, Nishimura K, Hirata-Saito A, Harada S, Arakawa M, Kozawa K, Noda M, Kimura H. Genetic analysis of the VP4/VP2 coding region in human rhinovirus species C in patients with acute respiratory infection in Japan. *J Med Microbiol.* 2013 Apr ; 62(Pt 4):610-7.

2. 学会発表等

なし

H. 知的財産権の出願・登録状況

なし

研究成果の刊行に関する一覧表

雑誌

発表者氏名	論文タイトル名	発表誌名	巻・号	ページ	出版年
Fumihiko Takeuchi, Tsuyoshi Sekizuka, Akifumi Yamashita, Yumiko Ogasawara, Katsumi Mizuta, and Makoto Kuroda.	MePIC, Metagenomic Pathogen Identification for Clinical Specimens.	Jpn. J. Infect. Dis.	67 (1)	62-65.	2014
Senchi K, Matsunaga S, Hasegawa H, Kimura H, Ryo A.	Development of oligomannose-coated liposome-based nasal vaccine against human parainfluenza virus type 3.	Front in Microbiol.	4:	346	2013.
Kimura H, Yoshizumi M, Ishii H, Oishi K, Ryo A.	Cytokine production and signaling pathways in respiratory virus infection.	Front Microbiol	4	276	2013
Kushibuchi I, Kobayashi M, Kusaka T, Tsukagoshi H, Ryo A, Yoshida A, Ishii H, Saraya T, Kurai D, Yamamoto N, Kanou K, Saitoh M, Noda M, Kuroda M, Morita Y, Kozawa K, Oishi K, Tashiro M, Kimura H.	Molecular evolution of attachment glycoprotein (G) gene in human respiratory syncytial virus detected in Japan 2008-2011.	Infect Genet Evol	18	168-73	2013
Ishigami T, Abe K, Aoki I, Minegishi S, Ryo A, Matsunaga S, Matsuoka K, Takeda H, Sawasaki T, Umemura S, Endo Y.	Anti-interleukin-5 and multiple autoantibodies are associated with human atherosclerotic diseases and serum interleukin-5 levels.	FASEB J.	27(9)	3437-45	2013

Kiyota N, Kushibuchi I, Kobayashi M, Tsukagoshi H, Ryo A, Nishimura K, Hirata-Saito A, Harada S, Arakawa M, Kozawa K, Noda M, Kimura H.	Genetic analysis of the VP4/VP2 coding region in human rhinovirus species C in patients with acute respiratory infection in Japan.	J Med Microbiol.	62(Pt 4)	610-7	2013
Kobayashi M, Takayama I, Kageyama T, Tsukagoshi H, Saitoh M, Ishioka T, Yokota Y, Kimura H, Tashiro M, Kozawa K.	A new reassortant swine influenza A (H1N2) virus derived from A (H1N1) pdm09 virus isolated from swine.	Emerg Infect Dis.	19(12):	1972-4	2013
Yamazaki M, Sugai K, Kobayashi Y, Kaburagi Y, Murashita K, Saito N, Niino H, Imagawa T, Tsukagoshi H, Kimura H.	A child case of hypocomplementemic urticarial vasculitis due to Coxsackievirus type A9.	J Med Microbiol Case	1(1):	1-5	2014
Miyaji Y, Kobayashi M, Sugai K, Tsukagoshi H, Niwa S, Fujitsuka-Nozawa A, Noda M, Kozawa K, Yamazaki F, Mori M, Yokota S, Kimura H.	Respiratory severity and virus profiles in Japanese children with acute respiratory illness.	Microbiol Immunol.	57(12):	811-821	2013
Kiyota N, Kobayashi M, Tsukagoshi H, Ryo A, Harada S, Kusaka T, Obuchi M, Shimojo N, Noda M, Kimura H.	Genetic analysis of human rhinovirus species A to C detected in patients with acute respiratory infection in Kumamoto prefecture, Japan 2011–2012,	Infect Mol Evol.	21:	90-102	2013

Abe M, Tahara M, Yamaguchi H, Kanou K, Shirato K, Kawase M, Noda M, Kimura H, Matsuyama S, Fukuhara H, Mizuta K, Maenaka K, Ami Y, Esumi M, Kato A, Takeda M.	TMPRSS2 is an activating protease for respiratory parainfluenza viruses.	J Virol.	87(21):	11930-11935	2013
Matsuda S, Nakamura M, Hirano E, Kiyota N, Omura T, Suzuki Y, Noda M, Kimura H.	Characteristics of human metapneumovirus infection prevailing in hospital wards housing patients with severe disabilities.	Jpn J Infect Dis.	66(3):	195-200	2013
Tsukagoshi H, Yokoi H, Kobayashi M, Kushibuchi I, Okamoto-Nakagawa R, Yoshida A, Morita A, Noda M, Yamamoto N, Sugai K, Oishi K, Kozawa K, Kuroda M, Shirabe K, Kimura H.	Genetic analysis of attachment glycoprotein (G) gene in new genotype ON1 of human respiratory syncytial virus detected in Japan.	Microbiol Immunol.	57(9):	655-659	2013
Ishioka T, Yamada Y, Kimura H, Yoshizumi M, Tsukagoshi H, Kozawa K, Maruyama K, Hayashi Y, Kato M.	Elevated macrophage inflammatory protein 1α and interleukin-17 production in an experimental asthma model infected with respiratory syncytial virus.	Int Arch Allergy Immunol.	161(2):	129-137	2013
Mizuta K, Yamakawa T, Nagasawa H, Itagaki T, Katsushima F, Katsushima Y, Shimizue Y, Ito S, Aokia Y, Ikeda T, Abiko C, Kuroda M, Noda M, Kimura H, Ahiko T.	Epidemic myalgia associated with human parechovirus type 3 infection among adults occurs during an outbreak among children: Findings from Yamagata, Japan, in 2011.	J Clin Virol.	58(1):	188-193	2013

Saraya T, Mikoshiba M, Kamiyama H, Yoshizumi M, Tsuchida S, Tsukagoshi H, Ishioka T, Terada M, Tanabe E, Tomioka C, Ishii H, Kimura H, Kozawa K, Shiohara T, Takizawa T, Goto T.	Evidence for reactivation of human herpes virus 6 in generalized lymphadenopathy in a patient with drug induced hypersensitivity syndrome.	J Clin Microbiol.	51(6):	1979-1982	2013
Mizuta K, Abiko C, Aoki Y, Ikeda T, Matsuzaki Y, Itagaki T, Katsushima F, Katsushima Y, Noda M, Kimura H, Ahiko	T. Seasonal patterns of respiratory syncytial virus, influenza A virus, human metapneumovirus, and parainfluenza virus type 3 infections based on virus isolation data between 2004 and 2011 in Yamagata, Japan.	Jpn J Infect Dis.	66(2):	140-145	2013
Kiyota N, Kushibuchi I, Kobayashi M, Tsukagoshi H, Ryo A, Nishimura K, Hirata-Saito A, Harada S, Arakawa M, Kozawa K, Noda M, Kimura H.	Genetic analysis of VP4/VP2 coding region in human rhinovirus species C detected from the patients with acute respiratory infection in Japan.	J Med Microbiol.	62,	610-617	2013
Seki E, Yoshizumi M, Tanaka R, Ryo A, Ishioka T, Tsukagoshi H, Kozawa K, Okayama Y, Goya T, Kimura H.	Cytokine profiles, signaling pathways, and effects of fluticasone propionate in respiratory syncytial virus-infected human fetal lung fibroblasts.	Cell Biol Int.	in press.		2014
Nishi M, Sakai Y, Akutsu H, Nagashima Y, Quinn G, Masui S, Kimura H, Perrem K, Umezawa A, Yamamoto N,Lee SW, Ryo A.	Induction of cells with cancer stem-cell properties from non-tumorigenic human mammary epithelial cells by defined reprogramming factors.	Oncogene	in press.		2014

Abiko C, Mizuta K, Aoki Y, Ikeda T, Itagaki T, Noda M, Kimura H, Ahiko T.	An outbreak of parainfluenza virus type 4 infections among children with acute respiratory infections, in the 2011-12 winter season in Yamagata, Japan.	Jpn J Infect Dis.	66(1):	76-78	2013
Nakamura M, Hirano E, Ishiguro F, Mizuta K, Noda M, Tanaka R, Tsukagoshi H, Kimura H.	Molecular epidemiology of human metapneumovirus from 2005 to 2011 in Fukui, Japan.	Jpn J Infect Dis.	66(1):	56-59	2013
Kobayashi M, Tsukagoshi H, Ishioka T, Mizuta K, Noda M, Morita Y, Ryo A, Kozawa K, Kimura H.	Seroepidemiology of saffold cardiovirus (SAFV) genotype 3 in Japan.	J Infect.	66(2):	191-193	2013
Okada S, Hasegawa S, Hasegawa H, Ainai A, Atsuta R, Ikemoto K, Sasaki K, Toda S, Shirabe K, Takahara M, Harada S, Morishima T, Ichiyama T.	Analysis of bronchoalveolar lavage fluid in a mouse model of bronchial asthma and H1N1 2009 infection.	Cytokine.	63(2):	194-200.	2013

学会発表一覧表

発表者氏名	発表タイトル名	学会名	開催年月日	開催地
Miyaji Y, Kobayashi M, Sugai K, Tsu-kagoshi H, Niwa S, Fujitsuka-Nozawa A, Noda M, Kozawa K, Yamazaki F, Mori M, Yokota S, Kimura H.	Severity of respirato-ry signs and symptoms and virus profiles in Japanese children with acute respiratory illness.	Asia Pacific Congress of Asthma, Allergy, and Clinical Immunology (AP-CAACI)	Nov 14-17, 2013	Taipei, Taiwan.
Obuchi M, Hatasaki K, Tsubata S, Kaneda H, Shinozaki K, Tsuji T, Kasei M, Konishi M, Inasaki N, Obara-Nagoya M, Horimoto E, Sata T and Takizawa T.	Detection of Resistant Influenza A(H3N2) Virus in Children Treated with Neuraminidase Inhibitors Using a Next-Generation DNA Sequencer.	53rd Interscience Conference on Antimicrobial Agents and Chemotherapy.	2013 年 9 月	米国デンバー
小渕正次、畑崎喜芳、津幡眞一、金田 尚、篠崎健太郎、辻隆男、紺井正春、小西道雄、稲崎倫子、名古屋（小原）真弓、堀元栄詞、佐多徹太郎、滝澤剛則.	ノイラミニダーゼ阻害薬投与小児患者における薬剤耐性 A(H3N2)インフルエンザウイルスの検出.	第 61 回日本ウイルス学会学術集会	2013 年 11 月.	神戸
片野晴隆、佐藤由子、中島典子、福本瞳、鈴木忠樹、黒田誠、長谷川秀樹	病理検体からの不明病原体検出法の最先端ワークショップ 「感染病理学の新展開」	第 102 回 日本病理学会総会.	2013.4.	札幌
中島典子、片野晴隆	定量的 PCR によるウイルスの網羅的検出法と病理検体への応用 シンポジウム３ 病原体の新しい診断法	第 18 回日本神経感染症学会総会学術集会	2013 年 10 月	宮崎

Jpn. J. Infect. Dis., 67, 62-65, 2014

Short Communication

MePIC, Metagenomic Pathogen Identification for Clinical Specimens

Fumihiko Takeuchi[1], Tsuyoshi Sekizuka[1], Akifumi Yamashita[1], Yumiko Ogasawara[1], Katsumi Mizuta[2], and Makoto Kuroda[1]*

[1]*Pathogen Genomics Center, National Institute of Infectious Diseases, Tokyo 162-8640; and* [2]*Department of Microbiology, Yamagata Prefectural Institute of Public Health, Yamagata 990-0031, Japan*

(Received July 10, 2013. Accepted September 26, 2013)

SUMMARY: Next-generation DNA sequencing technologies have led to a new method of identifying the causative agents of infectious diseases. The analysis comprises three steps. First, DNA/RNA is extracted and extensively sequenced from a specimen that includes the pathogen, human tissue and commensal microorganisms. Second, the sequenced reads are matched with a database of known sequences, and the organisms from which the individual reads were derived are inferred. Last, the percentages of the organisms' genomic sequences in the specimen (i.e., the metagenome) are estimated, and the pathogen is identified. The first and last steps have become easy due to the development of benchtop sequencers and metagenomic software. To facilitate the middle step, which requires computational resources and skill, we developed a cloud-computing pipeline, MePIC: "Metagenomic Pathogen Identification for Clinical specimens." In the pipeline, unnecessary bases are trimmed off the reads, and human reads are removed. For the remaining reads, similar sequences are searched in the database of known nucleotide sequences. The search is drastically sped up by using a cloud-computing system. The webpage interface can be used easily by clinicians and epidemiologists. We believe that the use of the MePIC pipeline will promote metagenomic pathogen identification and improve the understanding of infectious diseases.

Next-generation DNA sequencing technologies have led to a new method of identifying the causative agent of infectious diseases in hospitalized patients and during outbreaks (1,2). By directly sequencing millions of DNA/RNA molecules in a specimen and matching the sequences to those in a database, pathogens can be inferred. The analysis comprises three steps. First, the nucleotide sequences of the specimen, which includes the pathogen, human tissue and commensal microorganisms, are read using a next-generation sequencer. Second, from bioinformatic processing of the reads, the organisms from which the individual reads were derived are inferred. Last, the percentages of the organisms' genomic sequences in the specimen are estimated, and the pathogen is identified. Although the first and last steps have become easy due to the development of benchtop sequencers and metagenomic software, the middle step still requires computational resources and bioinformatic skill. To facilitate the middle step, we developed a cloud-computing pipeline that is easy and fast.

The prototype of the pipeline has been used in our metagenomic search for pathogens in various clinical cases. We have reported metagenomic analyses of clinical specimens that successfully identified *Francisella tularensis* in an abscess as a pathogen (3), *Streptococcus* spp. in a lymph node as a possible causative candidate of Kawasaki disease (4) and heterogeneity of the 2009 pandemic influenza A virus (A/H1N1/2009) in the lung (5). In an outbreak of 22 adults with myalgia, the majority were infected with human parechovirus type 3, which typically causes disease in young children (6). In recent food poisoning outbreaks that were due to raw fish consumption, a flounder parasite *Kudoa septempunctata* was discovered as the causative agent (7). In all cases, the pathogen identification was primarily due to metagenomic analyses using next-generation sequencers.

In the workflow of metagenomic pathogen identification using MePIC, the first step is performed by the user. DNA and/or RNA is extracted from a specimen, such as sputum, feces, an abscess or blood, and a library of DNA/cDNA is prepared for sequencing. The library is sequenced using a benchtop next-generation sequencer, and the sequenced reads are uploaded to the MePIC pipeline via a secure internet connection. The pipeline accepts input files in FASTQ format, which is the standard for next-generation sequencing analysis. When using Illumina MiSeq sequencers (San Diego, Calif., USA).

In the second step, the uploaded reads are processed by the MePIC pipeline (Fig. 1). Unnecessary adapter sequences and low quality bases are trimmed off the reads using the fastq-mcf program in the ea-utils package (http://code.google.com/p/ea-utils/). Human-derived reads are detected through comparisons with the human genome using the BWA (8) program; the

* Corresponding author: Mailing address: Pathogen Genomics Center, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan. Tel: +81-3-5285-1111, Fax: +81-3-5285-1166, E-mail: makokuro@niid.go.jp

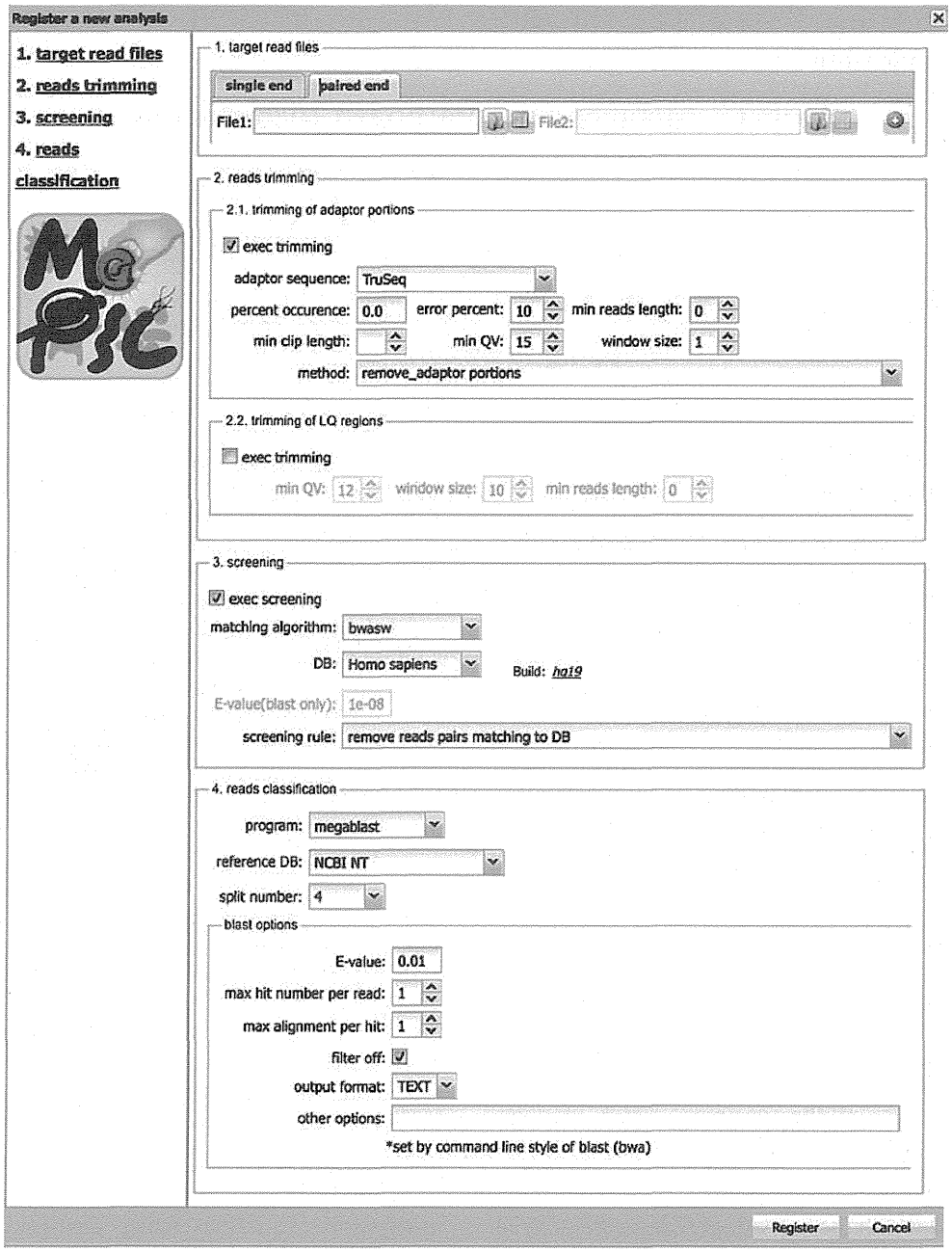

Fig. 1. Screenshot of the MePIC pipeline. In box 1, the user specifies the next-generation sequencer reads for upload. In box 2, details are set for trimming adaptor sequences and low quality bases from the reads. In box 3, criteria are set for the exclusion of human reads. In box 4, the user chooses the program for searching the database of known sequences.

number of human reads are counted, but the reads themselves are removed from the downstream analysis. For each of the remaining reads, similar sequences are searched in the database of all known nucleotide sequences (NCBI nt) using the MEGABLAST (9) or BWA program (10). Based on the information of the database sequences that match with the read, we can infer the gene (e.g., virulence gene) and organism (e.g., *Escherichia coli*) that the read is derived from.

The run time of the pipeline is primarily allotted to searching the database of known sequences. The required time can be drastically shortened by splitting the job and running in parallel using a cloud-computing system or local server. Respectively, it takes 10 h for one core of 2.67 GHz and 6 min for 100 cores to perform MEGABLAST search against the nt nucleotide database (as of year 2013) for one million reads of length 200 bp. The run time varies according to the sample source and condition. In blood samples, $>90\%$ of reads are derived from human and removed in the preprocessing step, and accordingly the time for database search of the remaining reads is reduced. Human derived reads are less in sputum samples (60%) or normal feces ($\sim 0\%$), which accordingly demands more database search time.

In the final step of the workflow, the user downloads the database search result to a local PC, including the reads annotated with the organism and gene function. To summarize the taxonomic and functional informa-

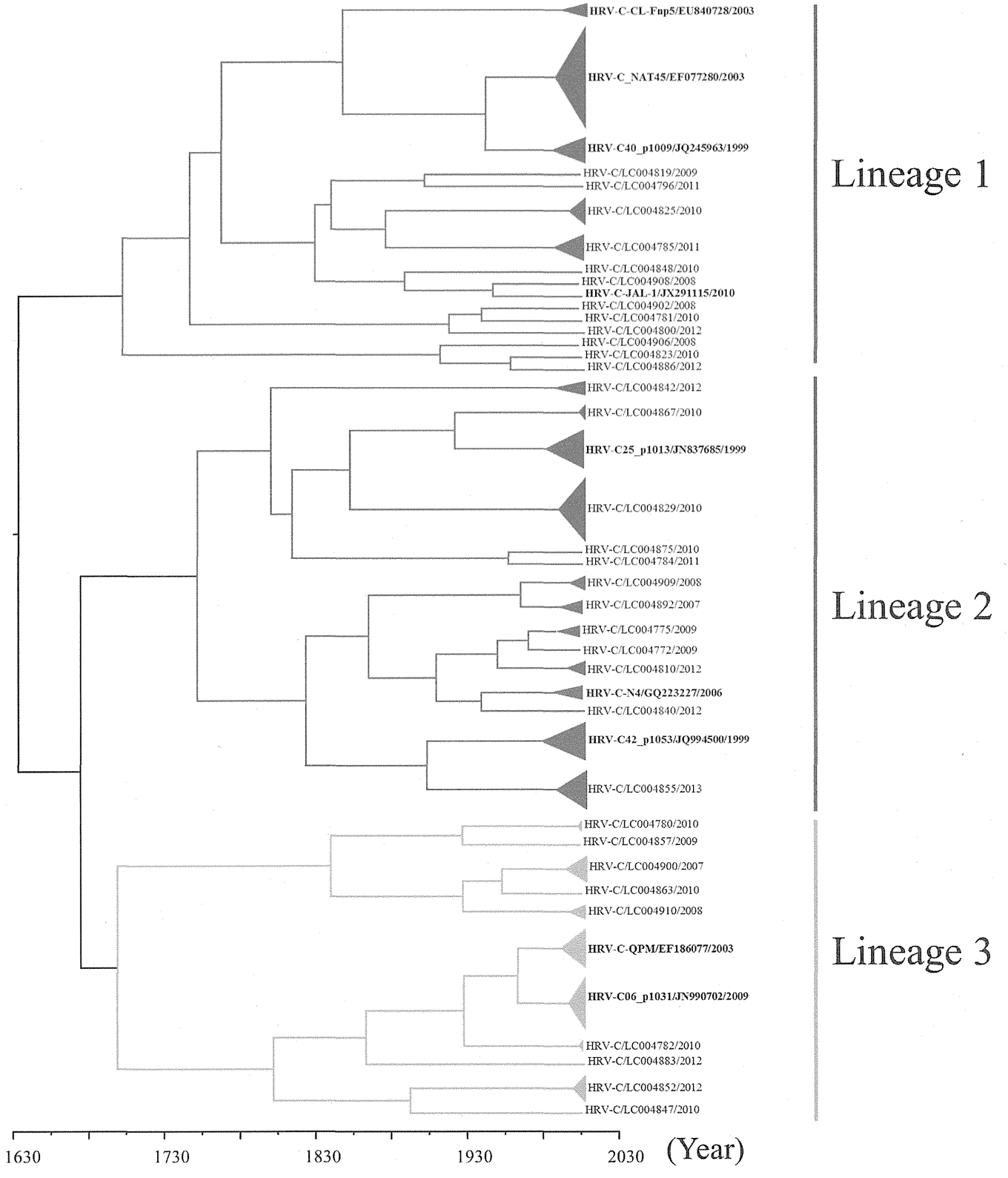

Figure 3 | Phylogenetic tree of the *VP3* gene in HRV-C constructed using the Bayesian MCMC method. A phylogenetic tree based on *VP3* gene sequences. In the tree, the triangle size expresses the number of strains, while only one representative strain is listed for each genotype. The strains are presented as follows: virus species/GenBank accession no./year. The GenBank accession numbers of the reference strains are indicated in bold letters. The scale bar represents the unit of time (year).

shown in Fig. 7 (a), an overall high similarity of the *VP1* coding region was found when compared to the *VP3* coding region. The minimum similarities of the *VP2, VP3* and *VP1* genes were approximately 70, 68 and 72%, respectively. Additionally, the similarities of the 5′-terminal *VP3* coding region and the 3′-terminal *VP1* coding region were low (approximately 70%), whereas the similarities of the 3′-terminal *VP3* coding region and the 5′-terminal *VP1* coding region were high (approximately 80%).

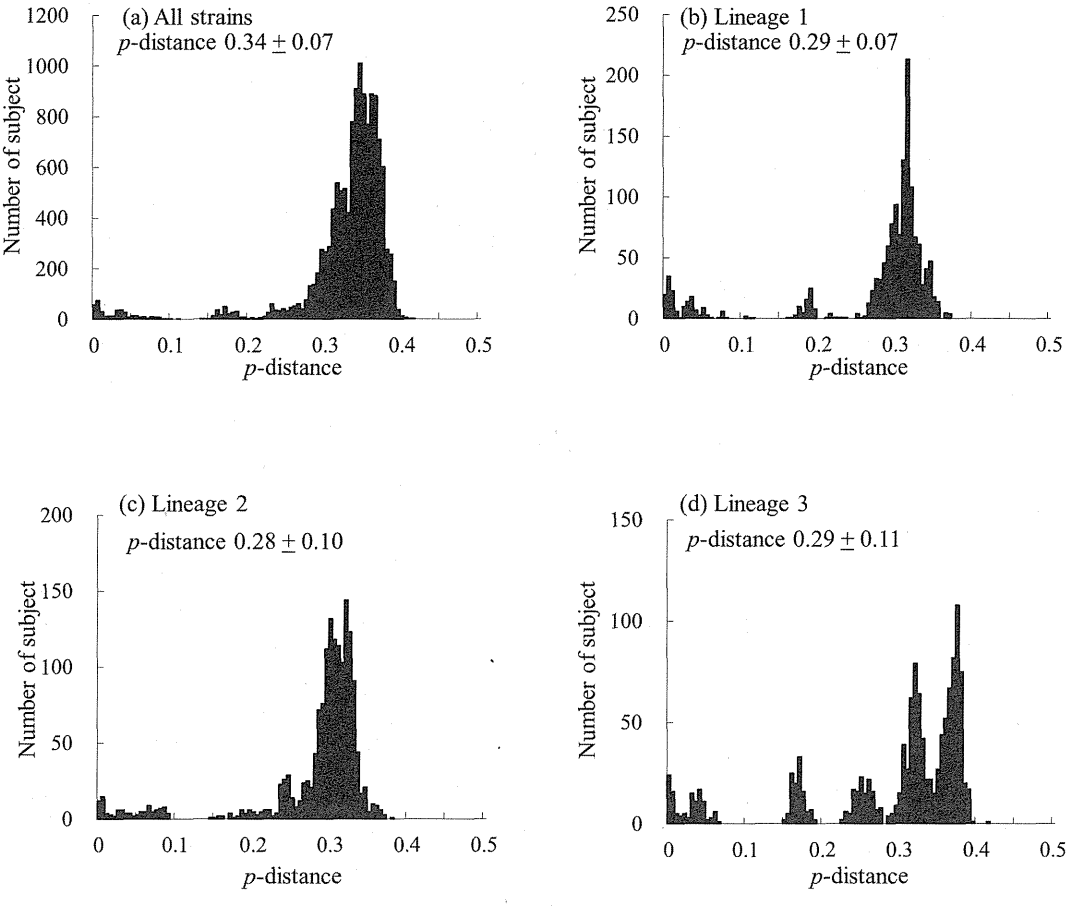

Figure 4 | The distribution of the pairwise interspecies distances based on the nucleotide sequences of the *VP1* gene. (a) The distribution of all strains. (b–d) The distributions of the pairwise distances for each lineage (lineages 1–3).

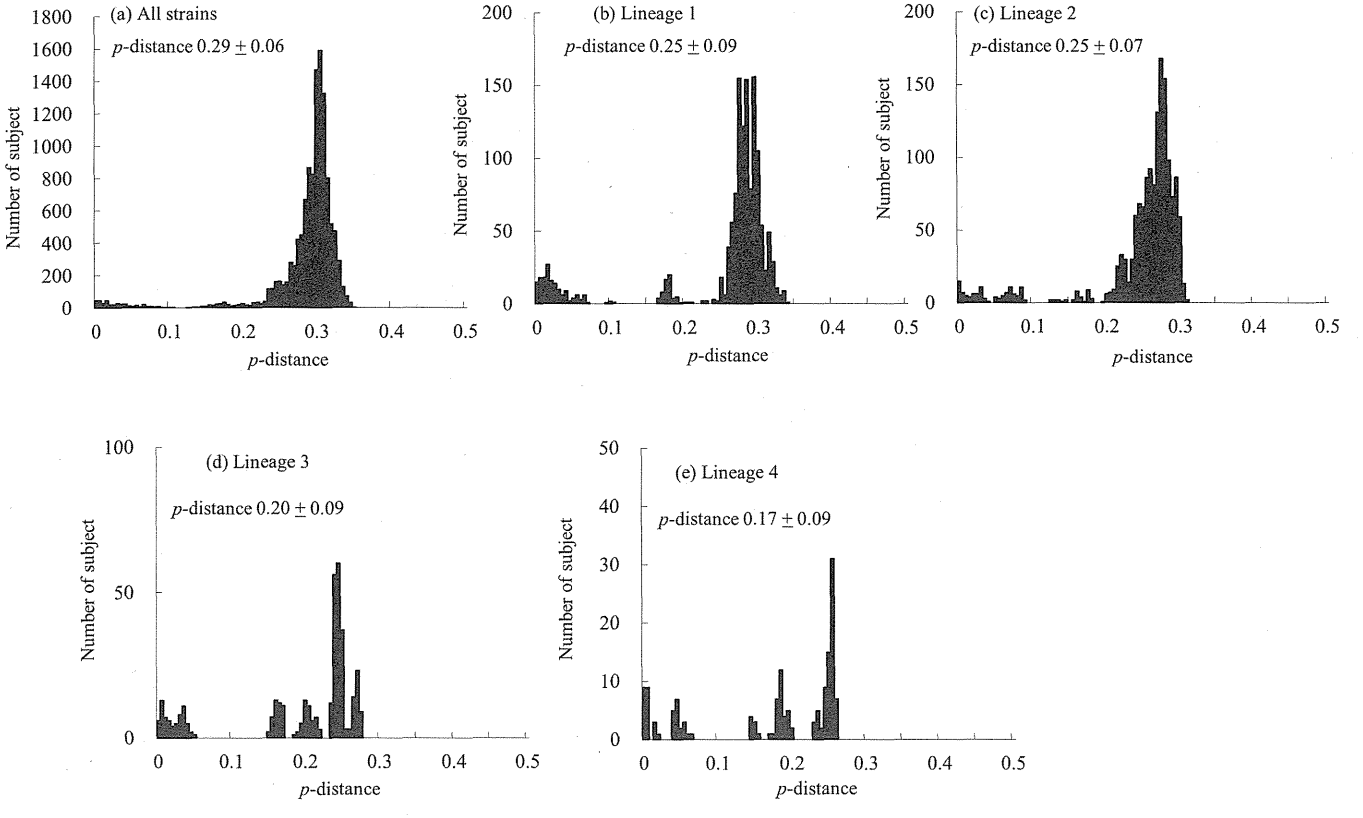

Figure 5 | The distribution of the pairwise interspecies distances based on the nucleotide sequences of the *VP2* gene. (a) The distribution of all strains. (b–e) The distributions of the pairwise distances for each lineage (lineages 1–4).

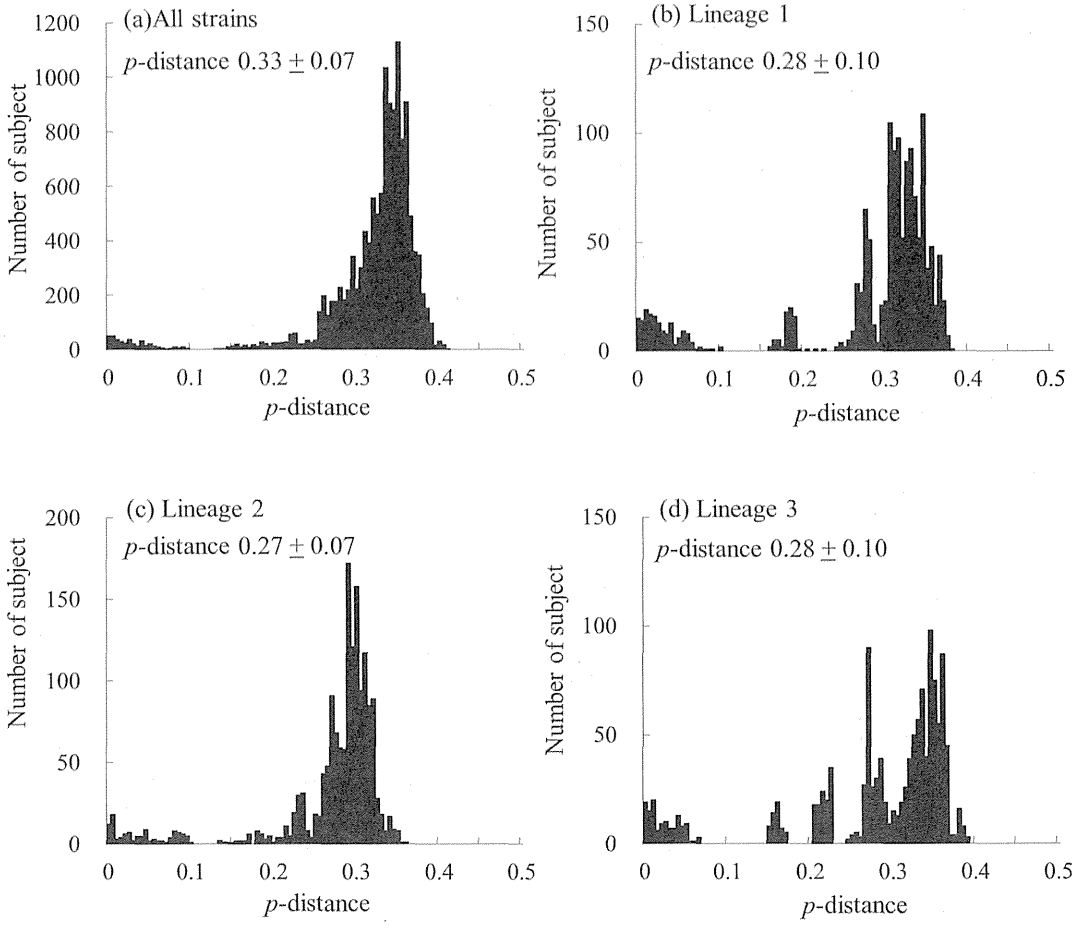

Figure 6 | The distribution of the pairwise interspecies distances based on the nucleotide sequences of the *VP3* gene. (a) The distribution of all strains. (b–d) The distributions of the pairwise distances for each lineage (lineages 1–3).

Additionally, we calculated the similarity of the deduced amino acid sequences of the *VP1* and *VP3* genes in the present strains and the prototype strain (HRV-QPM strain) (Fig. 7 (b)). The minimum similarities of the *VP2, VP3,* and *VP1* proteins were approximately 86, 83 and 87%, respectively. The Recombination Detection Program (RDP) found no evidence of a recombination event. The results suggested that high genetic divergence was found in the *VP3* coding region compared to the *VP2*, and *VP1* coding region.

Association between positive selection sites and the possible structures of the VP1, VP2 and VP3 proteins. Using SLAC, FEL, and IFEL methods, we estimated the positive selection sites in the VP1, VP2, and VP3 proteins in HRV-C. No positively selected sites were detected in any position by any method, while many sites under negative selection (>100) were found. Next, we constructed a molecular model of the complex containing the HRV-C VP1, VP2, and VP3 proteins. The model showed that these proteins are rich in loop structures that are primarily positioned on the exterior surface of the capsid complex (Figs. 8 (a) and 8 (b)). Our Shannon entropy data show that these exterior loops of VP1, VP2 and VP3 are highly variable compared to the interior of the capsid within the HRV-C population analyzed in this study (Figs. 8 (a) and 8 (b)).

Discussion

We studied the molecular evolution of the full length *VP1*, *VP2*, and *VP3* genes in HRV-C detected from ARI. Analysis of the full sequences of the major 3 viral protein genes in the current HRV-C strains was performed using NGS with an improved RT-PCR method. Time-scaled phylogenetic analysis with evolution rate for the genes was analyzed using the Bayesian MCMC method. Additionally, we constructed the VP1, VP2, and VP3 proteins using an *in silico* method. First, the phylogenetic trees based on the *VP1*, *VP2*, and *VP3* genes showed that the current HRV-C strains were classified into 3 or 4 major lineages, and these lineages were subdivided into many genotypes (>40). The most recent common ancestor (tMRCA) of all the strains based on the *VP1*, *VP2*, and *VP3* genes was found in the years 1652, 1125 and 1628, respectively. The evolution rates of both genes were fast. Similarity plot data showed high genetic divergence of the 5′-terminal *VP3* coding region. Moreover, no positively selected site was found in the VP1, VP2 and VP3 proteins. Additionally, no recombination of the *VP1*, *VP2*, and *VP3* genes was found in the studied strains. The exterior surfaces of the VP1, VP2, and VP3 proteins are rich in loops and are highly variable within the HRV-C population. The results suggested that the *VP1, VP2* and *VP3* genes, which encode major structural proteins of HRV-C, uniquely and rapidly evolved without positive selections.

Comprehensive molecular evolutionary and/or molecular epidemiologic studies of HRV-ABCs have been reported[13,14]. However, almost all studies were partially analyzed with regard to the *VP4/VP2* coding region of HRV[14,15]. The genes coding the VP proteins have many hypervariable regions; thus, it is difficult to design common primers for the amplification of the *VP1*, *VP2*, and *VP3* genes[12,16]. In addition, it may not be possible to isolate HRV-C using conventional methods at this time[2]. Thus, the antigenicity of HRV-C is still unknown. In the present study, we used an improved RT-PCR method with new primer sets and NGS, and we detected and analyzed the full length *VP1*, *VP2* and *VP3* genes with a high probability

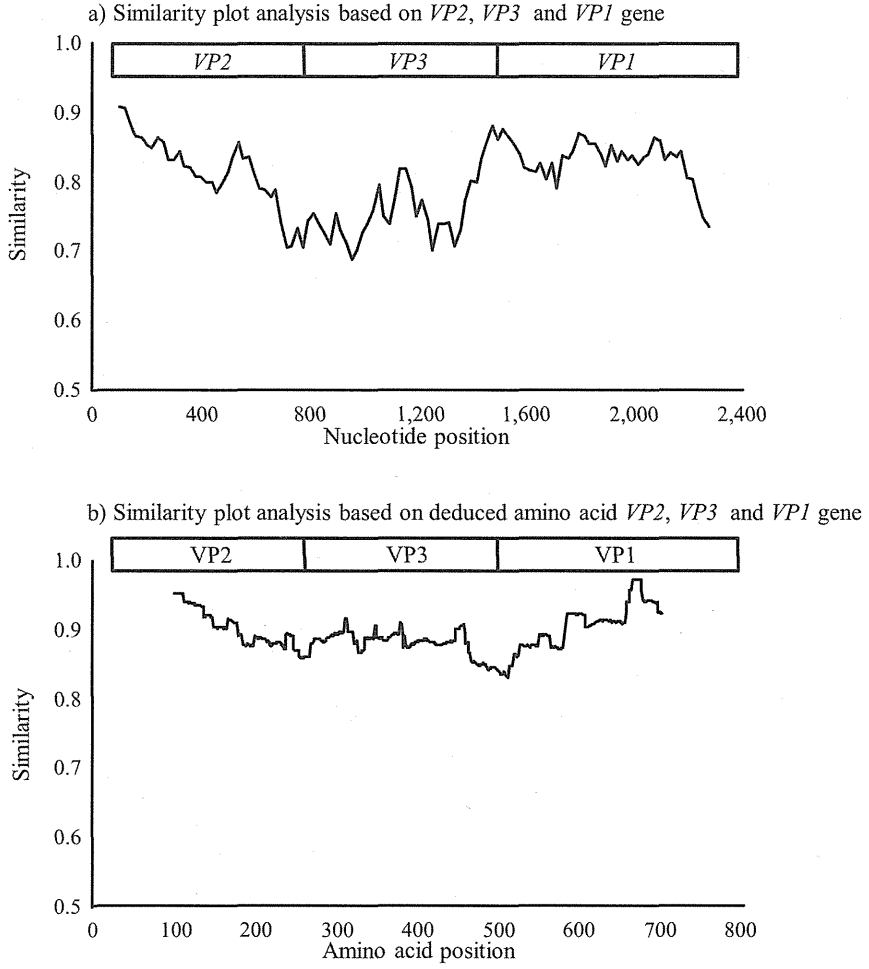

Figure 7 | Characterization of the *VP1*, *VP2* and *VP3* genes in HRV-C. (a) Nucleotide similarity to HRV-QPM (prototype strain) was determined using SimPlot analysis. (b) Amino acid similarity to HRV-QPM (prototype strain) was determined using SimPlot analysis. The plots indicate the percentage of similarity to a 50% consensus sequence from each species' polyprotein compared to HRV-QPM.

(>70%). To the best of our knowledge, this may be the first large study of the complete *VP1*, *VP2*, and *VP3* genes using many clinical strains.

Although HRV-C was recently discovered, it is thought that it may have a long history as a species. Indeed, Kiyota *et al.* showed that Japanese HRV-C strains could be dated back to the 1870s, according to the analysis of the *VP4/VP2* coding region[14]. The present strains dated back approximately 400 to 900 years (Fig. 1–3). Previous reports have suggested that HRV-C and HRV-A are frequently detected in patients with various ARIs[17]. Additionally, with regards

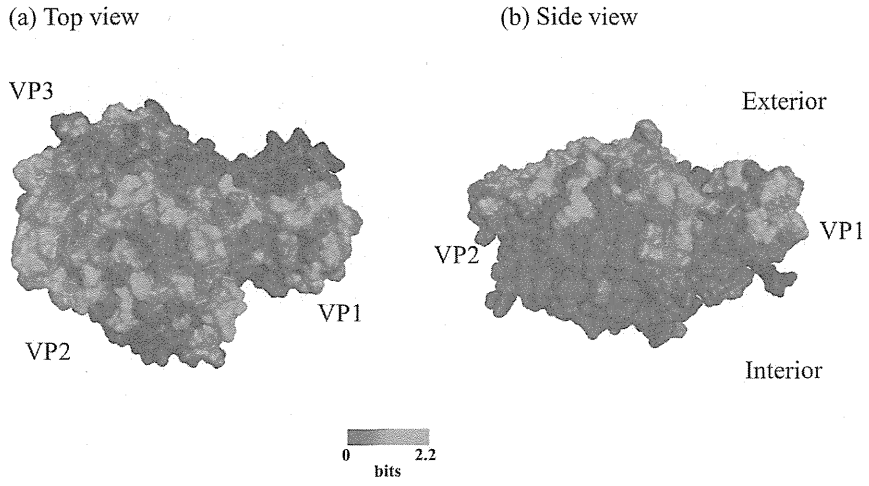

Figure 8 | Shannon entropy scores expressed on the structural models of HRV-C VP1, VP2, and VP3 proteins. The HRV-C VP1, VP2, and VP3 models were constructed by homology modeling as described in Materials and Methods and were superimposed on the VP1, VP2, and VP3 in those of the HRV2 capsid (PDB code: 3VDD). The entropy scores are expressed on the HRV-C VP1, VP2, and VP3 models. (a) Top view. (b) Side view.

to the analysis of the *VP4/VP2* coding region, both viruses exhibited large genetic divergence with many genotypes[15]. For example, Arakawa *et al.* showed that the *VP4/VP2* coding region in HRV-ABCs had over 0.3 divergence[17]. In the present study, greater than 0.3 divergence was found in the *VP1*, *VP2* and *VP3* genes. High genetic divergence in the 5′-terminal *VP3* coding region and the 3′-*VP1* coding region was found. The results from the partial analysis were compatible with our findings[12,18]. These results suggested that HRV-C might have a long history dating back at least 100 years; however, further studies are needed to confirm this.

The VP proteins of picornaviruses, which include HRV, play roles in their biology[7]. VP1, VP2, and VP3 protein are located at the surface of the viral capsid and are exposed to immune pressure, whereas VP4 is located inside the capsid[19]. For example, the VP1, VP2, and VP3 protein of many types of enteroviruses, such as EV71, is essential for the virus's ability to infect the host cells and acts as a protective protein in the viral shell[20]. Additionally, these proteins are recognized as major antigens in the host[20]. Indeed, the VP1 protein of many EVs, including HRV-A, is a major antigen[1]. However, the VP1, VP2, and VP3 proteins are major antigens for some types of HEVs[5]. It has been suggested that positive pressure in the host is associated with positive selection sites in major antigens[7]. Positive selection shows a survival advantage under the selective constraints that confront the viral population[7]. In the studied HRV-C strains, positive selection site was not found in the VP1, VP2 and VP3 proteins. Thus, HRV-C may be hardly affected under positive selection in our immune system.

Next, many negative selection sites (>100) were found in VP proteins in the present HRV-C strains. In general, negative selection plays an important role in maintaining the long-term stability of biological structures by removing deleterious mutations[7]. In the present study, many negative selection sites (>200) were found in both genes. Kiyota *et al.* showed that over 100 sites were found in the *VP4/VP2* coding region in HRV-C[14]. Thus, the negative selection sites in the VP1, VP2 and VP3 proteins may play the same roles as those in the *VP4/VP2* coding regions[21,22].

Our *in silico* structural analysis disclosed that the exterior surfaces of the VP1, VP2, and VP3 proteins are rich in loops, highly variable within the HRV-C population. It is conceivable that the exterior loops contain neutralization epitopes of HRV-C. In contrast, the interior regions of the VP1, VP2, and VP3 proteins were less diverse, suggesting the presence of functional and/or structural constraints on the diversity of this region. Some sites within these regions may be important for interactions with the infection receptor or the formation of a functional capsid complex structure. However, further studies may be needed to determine whether the VP1 protein is the major antigen in the infective HRV-C strains.

In conclusion, HRV-C was detected in various ARI patients, and the virus exhibited large genetic divergence with a uniquely rapid evolution. Additionally, these viruses have been agents of ARI for a lot longer than previously thought.

Methods

Samples and patients. Nasopharyngeal swabs were collected from 2,922 patients with ARI between November 2007 and March 2013. ARI patients were diagnosed mainly with upper respiratory infection (URI) or lower respiratory infection (LRI; bronchitis, bronchiolitis, and pneumonia). The samples were obtained by the local health authorities of the Fukui prefecture, Kumamoto prefecture, Tochigi prefecture, and Yokohama Medical Center for the surveillance of viral diseases in Japan. Informed consent was obtained from the patients or their guardian for the donation of the samples.

RNA extraction, RT-PCR and de novo sequencing by NGS. Viral RNA was extracted from 140 μL of supernatant using the QIAamp Viral RNA Mini Kit without carrier RNA (Qiagen, Valencia, CA). RT-PCR was performed using the PrimeScript® II High Fidelity One Step RT-PCR Kit (TaKaRa Bio, Otsu, Japan) and the primer pair HRV-C_546F: 5′- CTACTTTGGGTGTCCGTGTT -3′ and HRV-C_6410R: 5′-CCRTCATARTTDGTRTARTCAAA -3′. The PCR reactions are described in Suppl. Fig. S1. The NGS DNA library was prepared using a Nextera XT DNA sample prep kit (Illumina, San Diego, CA) with 96 indexing, followed by 200-mer paired-end *de novo* sequencing with MiSeq (Illumina). The obtained sequencing reads were assembled using the A5 assembler with the default parameters[23].

Phylogenetic analysis and estimation of the evolutionary rate using the Bayesian Markov chain Monte Carlo method. We aligned the nucleotide sequences of the *VP1, VP2* and *VP3* genes (positions; 2302-3126; 825 bp for HRV-QPM strain, positions 814-1602; 789 bp for HRV-QPM strain, positions 1603-2301; 699 bp for HRV-QPM strain) using CLUSTAL W [http://www.ddbj.nig.ac.jp/index-j.html]. To estimate the evolutionary rate and the time-scaled phylogeny, we used the Bayesian MCMC method in the BEAST package version 1.8.0[24]. The dataset was analyzed with a strict clock using the general time reversible with gamma-distributed rates across sites (GTR + I^-) substitution model[25,26] selected by the Kakusan4 program version 4.0 [http://www.fifthdimension.jp/products/kakusan/][27]. The MCMC chain was run for 50 million steps to achieve convergence, with sampling every 1000 steps. Convergence was assessed by the effective sample size (ESS) after a 10% burn-in using the Tracer program version 1.6 [http://tree.bio.ed.ac.uk/software/tracer]. Only parameters with an ESS above 200 were accepted. Uncertainty in the estimates was indicated by the 95% highest posterior density (HPD) intervals. The maximum clade credibility tree was obtained using the Tree Annotator program version 1.8.0, and the first 10% of the trees were removed as burn-in. The phylogenetic tree was viewed in the FigTree program version 1.5 [http://tree.bio.ed.ac.uk/software/figtree/].

Recombination analyses. Similarity plots showing the relationships between the aligned nucleotide sequences were generated using SimPlot, version 3.1 [http://sray.med.som.jhmi.edu/RaySoft/][28]. The level of nucleotide similarity in each sequence, with a window size of 200 nt and a step size of 20 nt, was calculated using the Kimura 2-parameter method, and similarity plot analyses based on the deduced amino acid sequences of the VP2, VP3 and VP1 proteins were performed with a window size of 100 aa. In this analysis, 1 aa was calculated using the Kimura 2-parameter method. The sequences were applied to RDP 3 [http://darwin.uvigo.es/rdp/rdp.html] to predict the recombination events using RDP, GENECONV, BootScan, Maxchi, Chimaera, SiSscan and 3Seq[29,30].

Selective pressure analysis and the calculation of pairwise distances. To obtain estimates of the positively and negatively selected sites among the present strains during each season, we calculated the synonymous (dS) and nonsynonymous (dN) rates at every codon in the alignment using Datamonkey [http://www.datamonkey.org/][31]. We used the following three different methods: single likelihood ancestor counting (SLAC), fixed effects likelihood (FEL), and internal fixed effects likelihood (IFEL). The SLAC and FEL methods were used to detect sites under selection at the external branches of the phylogenetic tree, while the IFEL method investigated sites along the internal branches. The SLAC method is best for large alignments but appears to underrate the substitution rate. Positively (dN > dS) and negatively (dN < dS) selected sites were determined by a *p*-value of <0.05 (SLAC, FEL, IFEL). Additionally, to assess the frequency distribution, we calculated the *p*-distance for the present strains, as previously described[10].

Molecular modeling of the HRV-C VP1, VP2, and VP3 proteins and analysis of amino acid diversity. Three-dimensional (3-D) models of the HRV-C VP1, VP2 and VP3 complex were constructed by homology modeling using 'MOE-Align' and 'MOE-Homology' in the Molecular Operating Environment (MOE) (Chemical Computing Group Inc., Quebec, Canada) as described for norovirus capsid protein modeling[32,33]. The X-ray crystal structures of HRV2 VP1 (PDB code: 3VDD), rhinovirus 14 capsid (PDB code:1R08) and human coxsackievirus VP3 (PDB code: 4GB3)[34] were used as the modeling templates for HRV-C VP1, VP2, and VP3 proteins, respectively, because these templates exhibited high scores with low E-values. The HRV-C VP1, VP2, and VP3 models were superimposed on VP1, VP2, and VP3 in the complex containing VP1, VP2, VP3, and VP4 of HRV2 strain (PDB code: 3VDD). Amino acid diversity at individual sites in the HRV-C sequences obtained in this study was analyzed with Shannon entropy scores as previously described[32].

Ethical approval. The study was approved by the National Institute of Infectious Disease Ethics Committee (No. 495), and the study was conducted in compliance with the principles of the Declaration of Helsinki.

Nucleotide sequence accession numbers. The sequences generated in this study have been assigned the GenBank accession numbers LC004772 to LC004910.

1. Pallansch, M. A. & Roos, R. Enteroviruses: Polioviruses, Coxsackieviruses, Echoviruses, and New Enteroviruses. Knipe, D. M., Howley, P. M., Griffin, D. E., Martin, M. A., Lamb, R. A., Roizman, B., Straus, S. E. (ed.), *Fields Virology* 839–893 Lippincott Willams & Wilkins, 2007.
2. Lee, W. M. *et al.* A diverse group of previously unrecognized human rhinoviruses are common causes of respiratory illnesses in infants. *PLoS One* **2**, e966 (2007).
3. Simmonds, P. *et al.* Proposals for the classification of human rhinovirus species C into genotypically assigned types. *J Gen Virol.* **91**, 2409–2419 (2010).
4. Lewis, J. K., Bothner, B., Smith, T. J. & Siuzdak, G. Antiviral agent blocks breathing of the common cold virus. *Proc Natl Acad Sci U S A.* **95**, 6774–6778 (1998).

5. Caro, V., Guillot, S., Delpeyroux, F. & Crainic, R. Molecular strategy for 'serotyping' of human enteroviruses. *J Gen Virol.* **82**, 79–91 (2001).
6. Sanjuán, R., Nebot, M. R., Chirico, N., Mansky, L. M. & Belshaw, R. Viral mutation rates. *J Virol.* **84**, 9733–9748 (2010).
7. Domingo, E. Virus evolution. Knipe, D. M., Howley, P. M., Griffin, D. E., Martin, M. A., Lamb, R. A., Roizman, B., Straus, S. E. (ed.), *Fields Virology.* 389–422 Lippincott Willams & Wilkins, 2007.
8. Thorne, J. L., Kishino, H. & Painter, I. S. Estimating the rate of evolution of the rate of molecular evolution. *Mol. Biol. Evol.* **15**, 1647–1657 (1998).
9. Kishibuchi, I. *et al.* Molecular evolution of attachment glycoprotein (G) gene in human respiratory syncytial virus detected in Japan 2008–2011. *Infect Genet Evol.* **18**, 168–173 (2013).
10. Mizuta, K. *et al.* Detailed genetic analysis of hemagglutinin-neuraminidase glycoprotein gene in human parainfluenza virus type 1 isolates from patients with acute respiratory infection between 2002 and 2009 in Yamagata prefecture, Japan. *Virol J.* **8**, 533 (2011).
11. Saitoh, M. *et al.* Molecular evolution of hemagglutinin (H) gene in measles virus genotypes D3, D5, D9, and H1. *PLoS One* **7**, e50660 (2012).
12. McIntyre, C. L., McWilliam Leitch, E. C., Savolainen-Kopra, C., Hovi, T. & Simmonds, P. Analysis of genetic diversity and sites of recombination in human rhinovirus species C. *J Virol.* **84**, 10297–10310 (2010).
13. McIntyre, C. L., Savolainen-Kopra, C., Hovi, T. & Simmonds, P. Recombination in the evolution of human rhinovirus genomes. *Arch Virol.* **158**, 1497–1515 (2013).
14. Kiyota, N. *et al.* Genetic analysis of the VP4/VP2 coding region in human rhinovirus species C in patients with acute respiratory infection in Japan. *J Med Microbiol.* **62**, 610–617 (2013).
15. Franco, D. *et al.* High genetic diversity and predominance of Rhinovirus A and C from Panamanian hospitalized children under five years with respiratory infections. *Virol J.* **9**, 257 (2012).
16. Tapparel, C. *et al.* Rhinovirus genome variation during chronic upper and lower respiratory tract infections. *PLoS One* **6**, e21163 (2011).
17. Arakawa, M. *et al.* Molecular epidemiological study of human rhinovirus species A, B and C from patients with acute respiratory illnesses in Japan. *J Med Microbiol.* **61**, 410–419 (2012).
18. Linsuwanon, P. *et al.* Complete coding sequence characterization and comparative analysis of the putative novel human rhinovirus (HRV) species C and B. *Virol J.* **8**, 5 (2011).
19. Racaniello, V. R. Picornaviridae: The viruses and their replication. Knipe, D. M., Howley, P. M., Griffin, D. E., Martin, M. A., Lamb, R. A., Roizman, B., Straus, S. E. (ed.), *Fields Virology* 795–838 Lippincott Willams & Wilkins, 2007.
20. Zhang, D., Lu, J. & Lu, J. Enterovirus 71 vaccine: close but still far. *Int J Infect Dis.* **14**, e739–e743 (2010).
21. Donker, N. C. & Kirkwood, C. D. Selection and evolutionary analysis in the nonstructural protein NSP2 of rotavirus A. *Infect Genet Evol.* **12**, 1355–1361 (2012).
22. Yoshida, A. *et al.* Molecular epidemiology of the attachment glycoprotein (G) gene in respiratory syncytial virus in children with acute respiratory infection in Japan in 2009/2010. *J Med Microbiol.* **61**, 820–829 (2012).
23. Tritt, A., Eisen, J. A., Facciotti, M. T. & Darling, A. E. An integrated pipeline for de novo assembly of microbial genomes. *PLoS One* **7**, e42304 (2012).
24. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol.* **7**, 214 (2007).
25. Drummond, A. J., Ho, S. Y., Phillips, M. J. & Rambaut, A. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**, e88 (2006).
26. Hasegawa, M., Kishino, H. & Yano, T. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol.* **22**, 160–174 (1985).
27. Tanabe, A. S. Kakusan4 and Aminosan: two programs for comparing nonpartitioned, proportional and separate models for combined molecular phylogenetic analyses of multilocus sequence data. *Mol Ecol Resour.* **11**, 914–921 (2011).
28. Zhang, Y. *et al.* Complete genome analysis of the C4 subgenotype strains of enterovirus 71: predominant recombinant C4 viruses persistently circulating in China for 14 years. *PLoS One* **8**, e56341 (2013).
29. McVean, G. A. *et al.* The fine-scale structure of recombination rate variation in the human genome. *Science* **304**, 581–584 (2004).
30. Martin, D. P. *et al.* RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* **26**, 2462–2463 (2010).
31. Pond, S. L. & Frost, S. D. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics.* **21**, 2531–2533 (2005).
32. Motomura, K. *et al.* Identification of monomorphic and divergent haplotypes in the 2006–2007 norovirus GII/4 epidemic population by genomewide tracing of evolutionary history. *J Virol.* **82**,11247–11262 (2008).
33. Motomura, K. *et al.* Divergent evolution of norovirus GII/4 by genome recombination from May 2006 to February 2009 in Japan. *J Virol.* **84**, 8085–8097 (2010).
34. Yoder, J. D., Cifuente, J. O., Pan, J., Bergelson, J. M. & Hafenstein, S. The crystal structure of a coxsackievirus B3-RD variant and a refined 9-angstrom cryo-electron microscopy reconstruction of the virus complexed with decay-accelerating factor (DAF) provide a new footprint of DAF on the virus surface. *J Virol.* **86**, 12571–12581 (2012).

Acknowledgments

This work was partly supported by a Grant-in-Aid from the Japan Society for the Promotion of Science and for Research on Emerging and Re-emerging Infectious Diseases from the Ministry of Health, Labour and Welfare, Japan (H25-Shinko-Ippan-015 and H22-Shinko-Ippan-011). We would like to thank Ms. Izumi Kushibuchi and Dr. Keiji Funatogawa (Tochigi Prefectural Institute of Public Health) and Ms. Eiko Hirano (Fukui Prefectural Institute of Public Health and Environmental Science) for helpful discussion.

Author contributions

M.K., A.R., M.T., K.O., H.I. and H.K. designed research; S.N., T.S., H.T., M.Y. and M.N. performed research; H.S., N.K., K.S. and M.O. contributed analytic tools, S.N., T.S., H.T., H.S., N.K., K.K., K.S. and T.K. analyzed data; M.K., N.S., S.H. and H.K. wrote the paper.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Kuroda, M. *et al.* Molecular evolution of the *VP1, VP2,* and *VP3* genes in human rhinovirus species C. *Sci. Rep.* **5**, 8185; DOI:10.1038/srep08185 (2015).

Infection, Genetics and Evolution 28 (2014) 183–191

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Molecular evolution of human respiratory syncytial virus attachment glycoprotein (*G*) gene of new genotype ON1 and ancestor NA1 ☆

Eiko Hirano [a], Miho Kobayashi [b], Hiroyuki Tsukagoshi [b], Lay Myint Yoshida [c], Makoto Kuroda [d], Masahiro Noda [e], Taisei Ishioka [e], Kunihisa Kozawa [b], Haruyuki Ishii [f], Ayako Yoshida [g], Kazunori Oishi [e], Akihide Ryo [h,*], Hirokazu Kimura [e,h,*]

[a] *Fukui Prefectural Institute of Public Health and Environmental Science, 39-4 Harame-cho, Fukui-shi, Fukui 910-8851, Japan*
[b] *Gunma Prefectural Institute of Public Health and Environmental Sciences, 378 Kamioki-machi, Maebashi-shi, Gunma 371-0052, Japan*
[c] *Department of Pediatric Infectious Diseases, Institute of Tropical Medicine, Nagasaki University, 1-12-4 Sakamoto, Nagasaki-shi, Nagasaki 852-8523, Japan*
[d] *Pathogen Genomics Center, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan*
[e] *Infectious Disease Surveillance Center, National Institute of Infectious Diseases, 4-7-1 Gakuen, Musashimurayama-shi, Tokyo 208-0011, Japan*
[f] *Department of Respiratory Medicine, Kyorin University, School of Medicine, 6-20-2 Shinkawa, Mitaka-shi, Tokyo 181-8611, Japan*
[g] *Aomori Prefectural Institute of Public Health and Environment, 1-1-1 Higashitsukurimichi, Aomori-shi, Aomori 030-8566, Japan*
[h] *Department of Microbiology and Molecular Biodefense Research, Yokohama City University, Graduate School of Medicine, 3-9 Fukuura, Kanagawa-ku, Yokohama-shi, Kanagawa 236-0004, Japan*

ARTICLE INFO

Article history:
Received 11 August 2014
Received in revised form 22 September 2014
Accepted 23 September 2014
Available online 2 October 2014

Keywords:
HRSV
Genotype ON1
Molecular evolution
Attachment glycoprotein (*G*) gene

ABSTRACT

We conducted a comprehensive genetic analysis of the C-terminal 3rd hypervariable region of the attachment glycoprotein (*G*) gene in human respiratory syncytial virus subgroup A (HRSV-A) genotype ON1 (93 strains) and ancestor NA1 (125 strains). Genotype ON1 contains a unique mutation of a 72 nucleotide tandem repeat insertion (corresponding to 24 amino acids) in the hypervariable region. The Bayesian Markov chain Monte Carlo (MCMC) method was used to conduct phylogenetic analysis and a time scale for evolution. We also calculated pairwise distances (*p*-distances) and estimated the selective pressure. Phylogenetic analysis showed that the analyzed ON1 and NA1 strains formed 4 lineages. A strain belonging to lineage 4 of ON1 showed wide genetic divergence (*p*-distance, 0.072), which suggests that it might be a candidate new genotype, namely ON2. The emergence of genotype NA1 was estimated to have occurred in 2000 (95% of highest probability density, HPD; 1997–2002) and that of genotype ON1 in 2005 (95% HPD; 2000–2010) based on the time-scaled phylogenetic tree. The evolutionary rate of genotype ON1 was higher than that of ancestral genotype NA1 (6.03×10^{-3} vs. 4.61×10^{-3} substitutions/site/year, $p < 0.05$). Some positive and many negative selection sites were found in both ON1 and NA1 strains. The results suggested that the new genotype ON1 is rapidly evolving with antigenic changes, leading to epidemics of HRSV infection in various countries.

1. Introduction

Human respiratory syncytial virus (HRSV) of genus *Pneumovirus* in family *Paramyxoviridae* is a major causative agent of acute lower respiratory infections. Specifically, primary HRSV infections may be responsible for about 50–90% of bronchitis, bronchiolitis, and pneumonia cases in children under 2 years of age (Leung et al., 2005; Shay et al., 1999; Yorita et al., 2007). Moreover, HRSV reinfections can occur throughout life and may result in bronchitis and pneumonia in elderly people (Lee et al., 2013). Thus, the impact of the disease burden of HRSV infection may be comparable to that of influenza (Lee et al., 2013; Weiss and McMichael, 2004).

Abbreviations: FEL, fixed effects likelihood; HPD, highest probability density; HRSV, human respiratory syncytial virus; IFEL, internal fixed effects likelihood; MCMC, Markov chain Monte Carlo; NJ, neighbor joining; *p*-distance, pairwise distance; SD, standard deviation; SLAC, single likelihood ancestor counting.

☆ The nucleotide sequences obtained in this study have been assigned Genbank: Nucleotide AB978289–AB978368

* Corresponding authors at: Infectious Disease Surveillance Center, National Institute of Infectious Diseases, 4-7-1 Gakuen, Musashimurayama-shi, Tokyo 208-0011, Japan. Tel.: +81 42 561 0771; fax: +81 42 565 3315 (H. Kimura). Department of Microbiology and Molecular Biodefense Research, Yokohama City University, Graduate School of Medicine, 3-9 Fukuura, Kanagawa-ku, Yokohama-shi, Kanagawa 236-0004, Japan. Tel.: +81 45 787 2600; fax: +81 45 787 2851 (A. Ryo).

E-mail addresses: aryo@yokohama-cu.ac.jp (A. Ryo), kimhiro@nih.go.jp (H. Kimura).

http://dx.doi.org/10.1016/j.meegid.2014.09.030
1567-1348/

HRSV contains two major antigens, fusion protein (F protein) and attachment glycoprotein (G protein) (Collins and Crowe, 2006). F protein is conserved, while rapid evolution (mutation) may be seen in the C-terminal 3rd hypervariable region of the G protein (Melero et al., 1997). This variable region contains some epitopes that induce neutralizing antibodies (Palomo et al., 1991). Thus, the ability of HRSV to establish reinfections throughout life may be due to the evolution of G protein, similar to the reinfections that occur with influenza virus subtype A(H1N1) due to the evolution of hemagglutinin (*HA*) gene (Hall et al., 1991; Taubenberger and Kash, 2010). HRSV have been classified into two subgroups, HRSV-A and HRSV-B, by genetic analysis of *G* gene/G protein (Mufson et al., 1985). Furthermore, HRSV-A and -B have been subdivided into 11 genotypes (GA1-GA7, SAA1, NA1, NA2, and ON1) and 20 genotypes (GB1-4, BA1-10, SAB1-4, and URU1-2), respectively (Cui et al., 2013; Trento et al., 2006). Of them, HRSV-A genotype ON1 was initially detected in 2010 in Ontario, Canada (Eshaghi et al., 2012). The results of genetic analyses suggest that new genotype ON1 evolved from genotype NA1 (Eshaghi et al., 2012). Notably, a tandem replication (corresponding to 24 amino acid residues) of 72 nucleotides was found in the C-terminal 3rd hypervariable region of the *G* gene of ON1 strains (Eshaghi et al., 2012). Furthermore, this genotype is rapidly replacing other genotypes, such as NA1, in some countries (Agoti et al., 2014; Kim et al., 2014; Pierangeli et al., 2014; Tsukagoshi et al., 2013). Another HRSV-B genotype BA first emerged in 1999 in Buenos Aires, Argentina (Galiano et al., 2005) and rapidly spread to various countries resulting in the current prevalence of HRSV-B infections (Trento et al., 2010). A tandem duplication of 60 nucleotides (corresponding to 20 amino acid residues) in the *G* gene C-terminal 3rd hypervariable region was also found in this genotype. Previous reports have suggested that the ancestor of genotype BA was another genotype, namely GB3 of HRSV-B (Galiano et al., 2005). BA has been subdivided into 10 genotypes (BA1-10) during the last 15 years (Dapat et al., 2010; Trento et al., 2006). However, the molecular evolution of the *G* gene in ON1 is not precisely known.

In general, the evolution of the virus may be associated with nucleic acid type (DNA or RNA), genome structure, and genome size (Gago et al., 2009; Holmes, 2011; Sanjuán et al., 2010). To gain a better understanding of HRSV epidemics, it is essential to analyze the *G* gene, the major antigen coding gene. Therefore, we conducted detailed genetic analyses of the global molecular evolution of the *G* gene of new prevalent HRSV genotype ON1 and its ancestor, NA1.

2. Materials and methods

2.1. Clinical samples

We collected nasopharyngeal swabs from patients with acute respiratory infection after verbal informed consent was obtained from the patients or their guardians. Samples were obtained between January 2009 and December 2013 in Fukui Prefecture. Patients were mainly diagnosed as having bronchitis. The study protocol was approved by the Ethics Committee of the National Institute of Infectious Diseases (approval No. 417).

2.2. RNA extraction, RT-PCR, sequencing, and BLAST search

RNA extraction was performed using a QIAamp Viral RNA Mini kit (QIAGEN, Valencia, CA), according to the manufacturer's instructions. Primescript RT reagent kit (Takara Bio, Otsu, Japan) was used to synthesize cDNA. We performed PCR using Takara ExTaq (Takara Bio) to amplify a part of the *G* gene, as previously described (Peret et al., 1998). The primer sequences used for PCR were as follows: first PCR primer set, forward primer F1 (5′-CAA CTCCATTGTTATTTGGC-3′), reverse primers GPA (5′-GAAGTGTTCA ACTTTGTACC-3′) and GPB (5′-AAGATGATTACCATTTTGAAG-3′); second PCR primer set, forward primer F1, reverse primers GSA (5′-AACCACCACCAAGCCCACAA-3′) and GSB (5′-AAAACCAACCAT CAAACCCAC-3′). PCR was carried out under the following conditions of 95 °C for 5 min, 30 cycles at 94 °C for 1 min, 50 °C for 1 min, and 72 °C for 2 min, followed by a final extension at 72 °C for 7 min.

Amplicons were purified with MinElute PCR purification kit (QIAGEN), and cycle sequenced by BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) using the second PCR primer sets. The products were purified with BigDye XTerminator® Purification Kit (Applied Biosystems). Sequence analysis was performed by an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems).

We conducted a BLAST search for the segments and genotyped all strains of HRSV-A. BLAST analyses indicated that there were 22 and 58 HRSV strains of ON1 and NA1 in the present samples, respectively. Of them, we omitted the strains with 100% nucleotide identity. Thus, we used 3 strains of ON1 and 20 strains of NA1 in this study.

2.3. Other strains used in this study

To estimate the global evolution of HRSV-A *G* gene in genotypes ON1 and its ancestor NA1, we obtained a comprehensive collection of the target region (C-terminal 3rd hypervariable region of *G* gene) from GenBank and the sequences were added to the dataset of the present strains. After alignment of the *G* gene sequences, we omitted the strains with 100% nucleotide identity. As a result, we analyzed 93 strains of genotype ON1 and 125 strains of NA1. Detailed data of the strains are shown in Table S1.

2.4. Calculation of pairwise distances

The nucleotide sequences of segments of the *G* gene (positions 658–894, 237 bp for strain AUS/A2/61, Genbank: Nucleotide M11486; positions 673–984, 309 bp for strain ON67-1210A, Genbank: Nucleotide JN257693) were aligned. The frequency distribution of pairwise distances (*p*-distances) among the strains of genotypes ON1 and NA1 were calculated using MEGA 6.0 (Tamura et al., 2013).

2.5. Phylogenetic analyses by the Bayesian Markov chain Monte Carlo and neighbor joining methods

Phylogenetic analyses by the Markov chain Monte Carlo (MCMC) method were performed as previously described (Kushibuchi et al., 2013). Briefly, we used Kakusan4 (http://www.fifthdimension.jp/products/kakusan/) to select the nucleotide substitution model (Tanabe, 2011). The datasets were analyzed by BEAST package program v1.7.5 (Drummond and Rambaut, 2007) under an uncorrelated lognormal relaxed clock model or a strict clock model (Drummond et al., 2006). The MCMC chains were run to achieve convergence with sampling every 1000 steps. Convergence was confirmed using Tracer v1.6.0 (http://tree.bio.ed.ac.uk/software/tracer/). We accepted parameters with effective samples size above 200 after 10% burn-in. The maximum clade credibility tree was generated by Tree Annotator v 1.7.4 after removing the first 10% of trees as burn-in and the phylogenetic tree was viewed in FigTree v.1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/). The evolutionary rates of genotypes NA1 and ON1 were also calculated as described above. Detailed conditions are shown in Table 1. We also constructed a phylogenetic tree by the neighbor

Table 1
Conditions for the estimation of the evolutionary time scale.

Genotype	No. of strain	Substitution model[a]	Clock model	Length of chain
ON1	93	HKY85-Γ	Lognormal relaxed clock	30,000,000
NA1	125	HKY85-Γ	Lognormal relaxed clock	15,000,000
All strains	236	HKY85-Γ	Strict clock	40,000,000

[a] HKY85: Hasegawa, Kishino and Yano 1985 model.

joining (NJ) method based on the analyzed region (Saitou and Nei, 1987). Evolutionary distances were estimated using Kimura's two-parameter method (Kimura, 1980). The reliability of the tree was estimated using 1,000 bootstrap replications.

2.6. Selective pressure analysis

To evaluate the selective pressures on the partial *G* gene among the strains, positive selection sites in each genotype were estimated by Datamonkey (http://www.datamonkey.org/), as previously described (Pond and Frost, 2005a). The synonymous (*dS*) and nonsynonymous (*dN*) substitution rates at every codon were calculated using the following three methods: single likelihood ancestor counting (SLAC), fixed effects likelihood (FEL), and internal fixed effects likelihood (IFEL). The cut-off *p*-value was set at 0.1 (Kushibuchi et al., 2013).

2.7. Statistical analysis

Statistical analyses were performed using Welch's test by EZR v1.24 (Kanda, 2013). A *p*-value of <0.05 was considered to be statistically significant.

3. Results

3.1. Phylogenetic analyses of the global ON1 and NA1 strains by MCMC and NJ methods

We constructed global time-scaled phylogenetic trees by the MCMC method using strains detected in various countries, including Japan, which were comprehensively collected from GenBank (Fig. 1a and b). We also constructed a phylogenetic tree by the NJ method to obtain a clear presentation of the genetic distances (Fig. 2). First, the phylogenetic trees obtained by MCMC method estimated that genotype ON1 diverged from genotype NA1 in 2005, while NA1 diverged from GA2 in 2000 (Table 2). An ancestor of all present strains could be dated back to 1953 (Fig. 1a and Table 2). These NA1 strains could be classified into 4 lineages (lineages 1–4, Fig. 1a). The new genotype ON1 strains were also classified into 4 lineages (lineages 1–4) on the phylogenetic trees constructed by both the MCMC and NJ methods (Figs. 1a and b and 2). The ON1 strains may have derived from lineage 1 of genotype NA1 (Fig. 1a). The ON1 strains belonging to lineage 1 (45 strains) were the dominant strains detected in many countries, namely Germany, Italy, Canada, Croatia, Thailand, Japan, South Korea, Philippines, China, and South Africa. Strains of lineage 2 (21 strains) were detected in Germany, Italy, Japan, and South Korea. Lineage 3 strains (25 strains) were from Kenya, Germany, Japan, Italy, and Croatia. Notably, the 2 strains detected in Italy (1301-118RM: Genbank: Nucleotide KC858255, 1301-125RM: Genbank: Nucleotide KC858256) independently formed lineage 4.

Next, the estimated divergence times for each lineage of new genotype ON1 were April 2009 (lineage 1), February 2010 (lineage 2), April 2010 (lineage 3), and August 2010 (lineage 4) (Fig. 1b). In addition, the evolutionary rate of ON1 strains in the analyzed region was higher than that of NA1 (mean rate, 6.03×10^{-3} vs. 4.61×10^{-3} substitutions/site/year, $p < 0.05$) (Table 2). The results suggested that new genotype ON1 evolved rapidly and spread quickly throughout many countries.

Table 2
Evolutionary rates and branched years of the analyzed HRSV-A genotypes.

Genotype	Mean rate[a] (95% HPD)	Branched year (95% HPD)
ON1	6.03×10^{-3} ($3.43–9.10 \times 10^{-3}$)*	2005 (2000–2010)
NA1	4.61×10^{-3} ($3.33–5.98 \times 10^{-3}$)*	2000 (1997–2002)
All strains[b]	5.36×10^{-3} ($4.42–6.39 \times 10^{-3}$)	1953 (1949–1956)

* $p < 0.05$.
[a] Substitutions/site/year.
[b] All strains are in the phylogenetic tree in this study.

3.2. The p-distance values and phylogenetic locations of lineages 1 to 3 of the ON1 strains

The distributions of the *p*-distances are shown in Fig. 3a and b. First, the mean values of the *p*-distances of genotypes ON1 and NA1 were relatively short (less than 0.025) and no significant differences were found; however, the distribution patterns differed. Next, the *p*-distance values of the present ON1 strains belonging to lineages 1–3 were <0.062.

3.3. New genotype candidate strain ON2 belonging to lineage 4 based on p-distance values and phylogenetic analyses by MCMC and NJ methods

It has been proposed that the assignment of each HRSV genotype corresponds to a *p*-distance value of less than 0.07 (Venter et al., 2001). To clearly demonstrate the *p*-distances of the present ON1 strains, we constructed a phylogenetic tree by the NJ method (Fig. 2). In the present study, two strains of ON1 formed an independent cluster as lineage 4 on the phylogenetic trees (Figs. 1a and b and 2). Of them, one strain (1301-118RM) had a large *p*-distance value of 0.072 (this value was calculated by the prototype ON1 strain, ON67-1210A). The other strain, 1301-125RM, had a relatively large *p*-distance value of 0.065. These results implicated that strain 1301-118RM is a candidate new genotype, namely ON2.

3.4. Positive and negative selection sites in the present strains

We estimated the positive and negative selection sites in the C-terminal 3rd hypervariable region of *G* gene in the present strains (Tables 3 and 4). In the ON1 strains, some sites under positive selection were found (Table 3). Of them, 2 substitutions were genotype specific (Asn251Asp, Asn251Tyr, or Asn251Ser; and Tyr297His). In particular, an amino acid substitution (Tyr297His) was found in ON1 strains in the newly inserted sites of the region (72 nt duplication). In the NA1 strains, 3 genotype specific sites under positive selection (Thr253Ile or Thr253Lys; Pro276Leu or Pro276Ser; and Thr296Ser, Thr296Ile, or Thr296Ala) were found (Table 3). Many sites under negative selection were found in both NA1 and ON1 strains (Table 4). These results suggested that

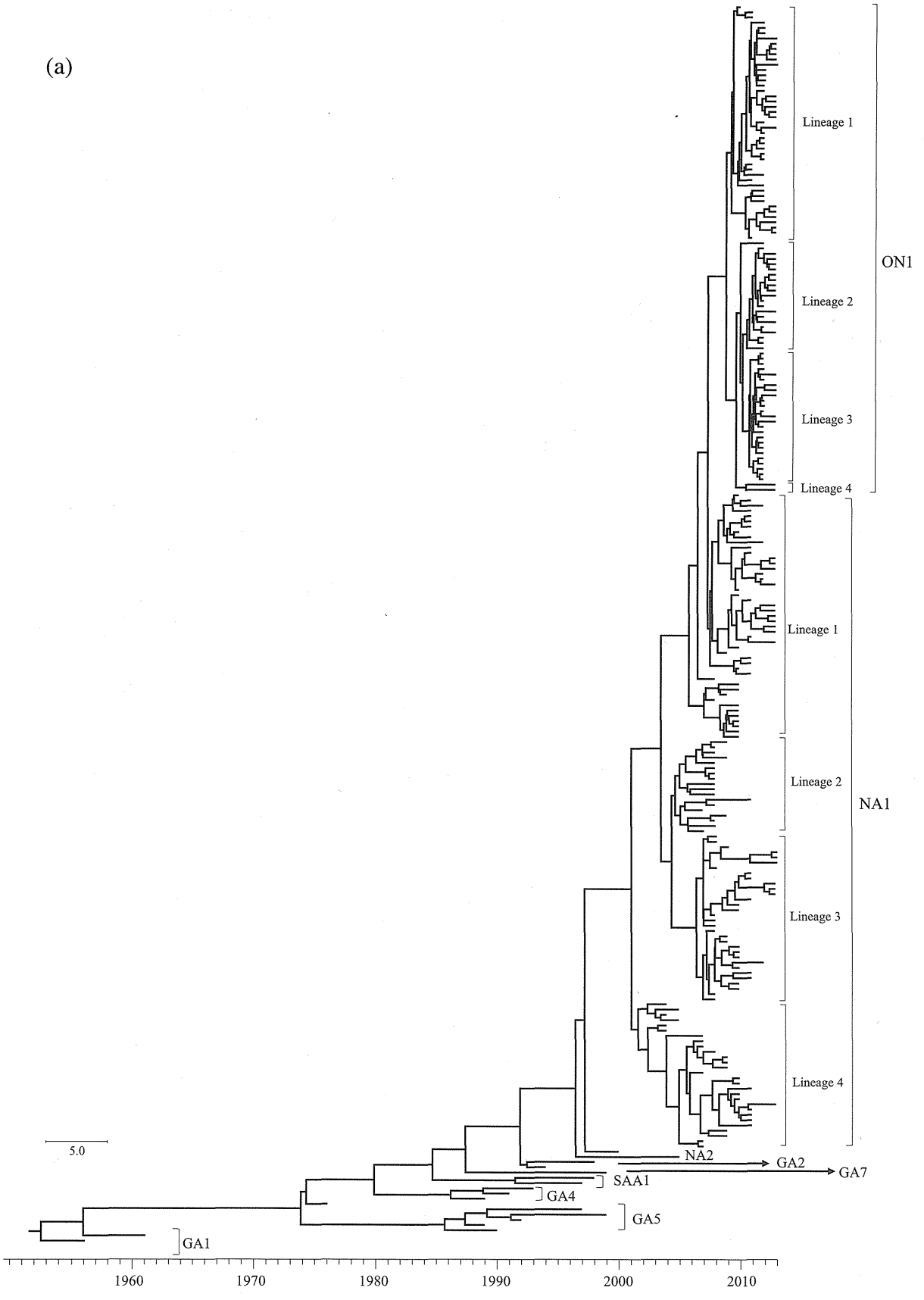

Fig. 1. Phylogenetic trees for *G* gene of HRSV-A (a) and the expanded genotype ON1 (b) constructed by the Bayesian Markov chain Monte Carlo (MCMC) method. Scale bars represent unit of time (year). Gray bars indicate 95% highest probability density (HPD) for the branched year.

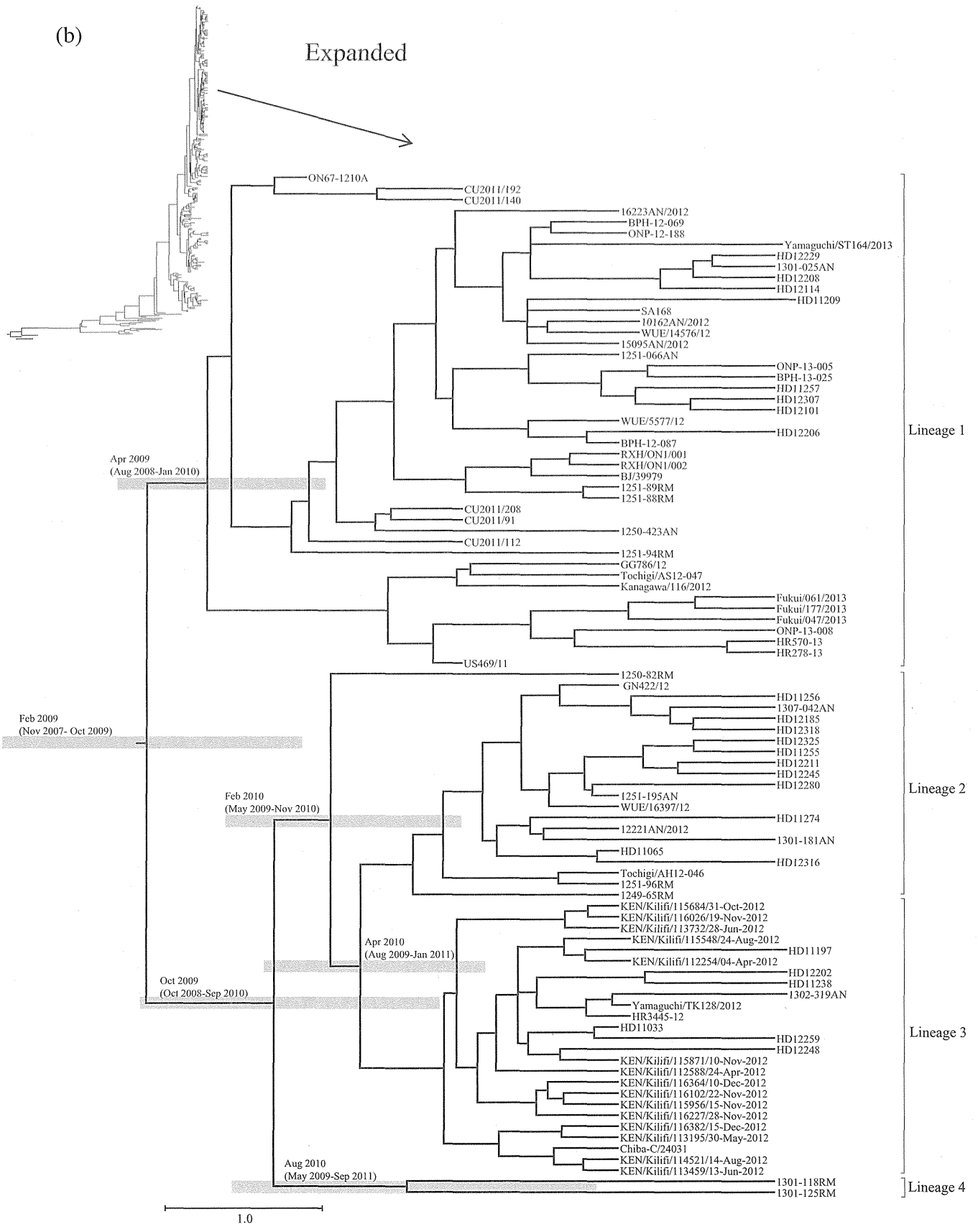

Fig. 1 (*continued*)

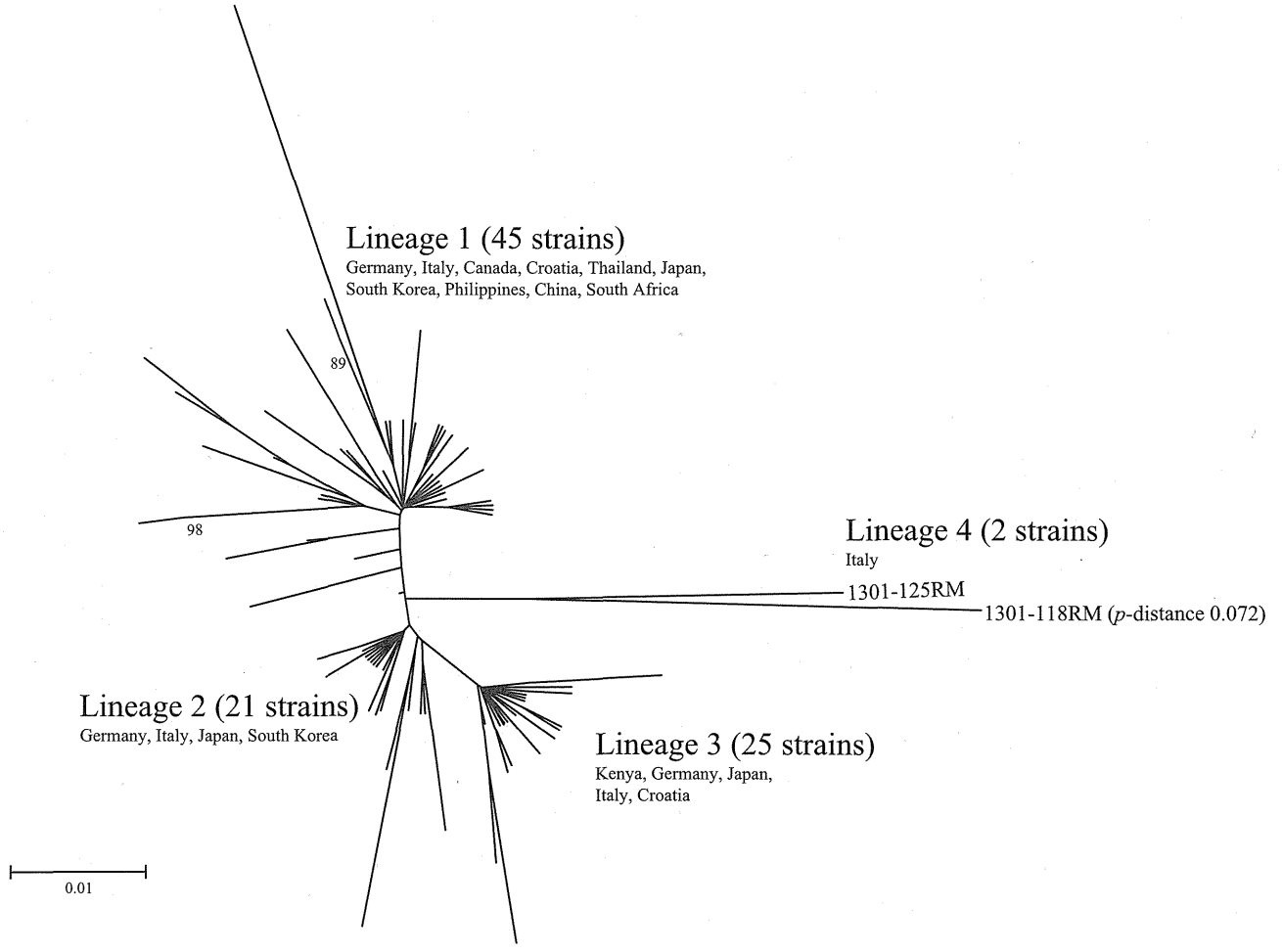

Fig. 2. Phylogenetic tree for *G* gene of HRSV-A genotype ON1constructed by the neighbor joining (NJ) method. Labels at the branch nodes show at least 70% bootstrap support. Scale bar indicates nucleotide substitutions per site.

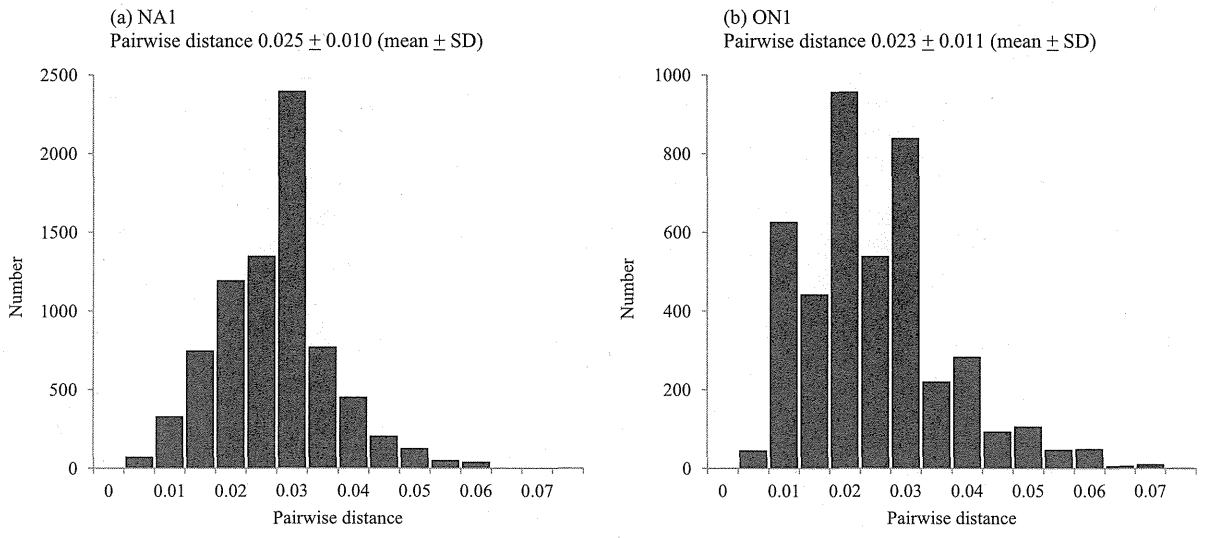

Fig. 3. Distribution of pairwise distances (*p*-distances) for HRSV-A genotype NA1 (a) and genotype ON1 (b) based on the nucleotide sequences of the *G* gene.

Table 3
Positive selection sites in C-terminal hypervariable region of *G* gene.

Genotype	Model	Positive selection site[a]	Mean *dN*/*dS*
ON1	SLAC	L274P, L274R	0.822
	FEL	N251D, N251Y, N251S, L274P, L274R, Y297H*	
	IFEL	S260N, Y273H, Y297H*	
NA1	SLAC	L274P	0.818
	FEL	T253I, T253K, S260N, S260I, N273Y, N273H, Y273N	
	IFEL	N273Y, N273H, Y273N, P276L, P276S, T296S, T296I, T296A	

p-Value < 0.1.
[a] Sites inside 72 duplication position are indicated asterisks.

Table 4
Negative selection sites in C-terminal hypervariable region of *G* gene.

Genotype	Model	Negative selection site[a]
ON1	SLAC	T227, P230, T231, T239, T245, S291*
	FEL	P222, E224, P230, T231, I243, T245, T259, S277, Q285, S291*, S307*
	IFEL	T231, S277, Q285, S291*
NA1	SLAC	T245, T268
	FEL	E232, T239, L248, G254, E257, L265, L266, T268, S270, T282, S287
	IFEL	L248, G254, T268, S270, S283

p-Value < 0.1.
[a] Sites inside 72 duplication position are indicated asterisks.

frequent amino acid substitutions have occurred in the analyzed region of genotypes NA1 and ON1.

4. Discussion

We studied the molecular evolution of the C-terminal 3rd hypervariable region of HRSV-A *G* gene in new genotype ON1 and its ancestor NA1. The phylogenetic trees with evolutionary time scales constructed by the MCMC method indicated that genotype ON1 diverged in 2005 from a lineage (lineage 1) of genotype NA1 (Fig. 1a). Four lineages of ON1 have emerged in various countries over some years and show a significantly rapid evolutionary rate in the analyzed region compared with genotype NA1 (Fig. 1a and Table 2). Furthermore, *p*-distance values and phylogenetic analyses suggested that a candidate new genotype (namely ON2), which likely derived from ON1, diverged in 2010. Some sites under positive selection and many under negative selection were found in both genotypes. The results suggested that new genotype ON1 is rapidly evolving with essential amino acid substitutions in the hypervariable region of the *G* gene.

Previous reports have deduced that new genotype ON1 of HRSV-A emerged in 2010 in Canada (Eshaghi et al., 2012) and may have been rapidly spreading and replacing another prevalent HRSV-A genotype, NA1, in many Asian, European, American, and African areas over a period of 3 years (Agoti et al., 2014; Auksornkitti et al., 2013; Cui et al., 2013; Eshaghi et al., 2012; Forcic et al., 2012; Kim et al., 2014; Pierangeli et al., 2014; Prifert et al., 2013; Tsukagoshi et al., 2013; Valley-Omar et al., 2013). This genotype has unique nucleotide insertions (72 nt duplication) in the C-terminal 3rd hypervariable region of HRSV-A *G* gene (Eshaghi et al., 2012). Genotype NA1, the likely ancestral strain of ON1, emerged in 2004 and became a prevalent type of HRSV-A infection in many areas in Asia, Europe, America, and Africa over a period of 10 years (Cui et al., 2013; de-Paris et al., 2014; Etemadi et al., 2013; Forcic et al., 2012; Khor et al., 2013; Kushibuchi et al., 2013; Pretorius et al., 2013; Rebuffo-Scheer et al., 2011;Shobugawa et al., 2009; Tran et al., 2013; Yamaguchi et al., 2011). A similar insertion has been confirmed in HRSV-B genotype BA (60 nt, corresponding to 20 amino acid insertions) (Trento et al., 2003). This genotype emerged in 1999 in Argentina (Galiano et al., 2005). The ancestral strain is thought to be genotype GB3 of HRSV-B, a prevalent type during the 1990s (Galiano et al., 2005). Furthermore, many divergent genotypes (BA1-10) have evolved from genotype BA over a period of 10 years, and are prevalent types of HRSV-B (Cui et al., 2013; Dapat et al., 2010; Trento et al., 2006). Previous reports have suggested that such nucleotide insertions in the *G* gene might be linked to changes in the antigenicity of the G protein (Trento et al., 2003). Thus, the mutations seen in HRSV-A genotype ON1 and HRSV-B genotype BA may lead to future epidemics of HRSV infections.

The analyzed region of ON1 strains showed a high evolutionary rate in comparison with ancestral genotype NA1 (Table 2). In addition, lineage 4 of ON1 showed wide genetic divergence in the phylogenetic tree (Fig. 2). Notably, the *p*-distance of one strain was calculated to be 0.072, based on the sequence of the prototype ON1 (ON67-1210A). This strain was detected in Rome, Italy in 2013. Previous reports proposed that the genetic distance (*p*-distance) range in the same genotype was <0.07 (Cui et al., 2013; Venter et al., 2001). When we apply the *p*-distance value, this strain may be a candidate new genotype (namely ON2) that evolved from ON1 (Figs. 1b and 2). Moreover, this strain formed a unique lineage (lineage 4) that may have emerged in August 2010 as estimated by the present phylogenetic tree (Fig. 1b). However, since we analyzed and evaluated only a part of the *G* gene, additional analysis of the strain, including whole genome analysis, may be needed.

The evolution of antigens of various respiratory viruses may be involved in the infectivity toward the host, including the ability to establish reinfections (Domingo, 2006). For example, the rapid evolution of *HA* gene in seasonal influenza viruses, such as subtypes A(H3N2) or A(H1N1), is closely related to the ability of influenza to reinfect the host (Taubenberger and Kash, 2010). Similarly, the evolution of HRSV *G* gene might be associated with the ability of HRSV to reinfect humans (Botosso et al., 2009; Collins and Melero, 2011). In the present study, we analyzed the evolution of the *G* gene in the prevalent genotypes of HRSV-A. Rapid rates of evolution were found in the analyzed region in both ON1 and NA1 strains, although the rate of ON1 was faster than that of NA1 (Table 2). The C-terminal 3rd hypervariable regions are known to be involved in the function of epitopes against neutralizing antibodies (Palomo et al., 1991). Thus, the high evolutionary rate of the analyzed region of HRSV may be associated with the ability of the virus to reinfect the host (Botosso et al., 2009; Collins and Melero, 2011).

It has been suggested that the evolution of major antigens of various respiratory viruses including HRSV is associated with selective pressure in the host (Botosso et al., 2009). Furthermore, negative selection may be associated with preventing deterioration of viral functions (Domingo, 2006). Thus, we analyzed sites under positive and negative selection in the analyzed regions. In both ON1 and NA1 strains, some sites under positive selection were

found, with two unique sites found in the ON1 strains (Asn251Asp, Asn251Tyr, or Asn251Ser; and Tyr297His). Of them, Tyr297His was located in the new tandem duplication regions of the genotype. In this study, we analyzed sites under positive selection using SLAC, IFL, and IFEL methods (Botosso et al., 2009; Kushibuchi et al., 2013). SLAC is the more conservative of the three methods and appropriate for large alignments (Pond and Frost, 2005b). However, the number of positively selected sites may be underestimated (Pond and Frost, 2005b). In contrast, FEL and IFEL methods take synonymous and nonsynonymous rate variations into account and may be efficiently parallelized (Pond and Frost, 2005b). Thus, we used the three different methods to obtain an accurate estimate of sites under positive selection in the present study (Botosso et al., 2009; Kushibuchi et al., 2013).

Furthermore, many sites under negative selection were found in both ON1 and NA1 strains (Table 4). In general, negative selection may act to prevent deterioration of various viruses (Domingo, 2006). For example, the sites under negative selection in neutralization epitopes of polioviruses may be involved in receptor recognition and in the formation of altered particles (Domingo et al., 1993). Although the roles of many sites under negative selection in HRSV G protein are not exactly known, it is possible that these amino acid substitutions are involved in preventing the deterioration of antigenic function (Domingo, 2006; Kushibuchi et al., 2013).

5. Conclusion

A prevalent new genotype, ON1 of HRSV-A, with some positively selected amino acid substitutions emerged during a few years of rapid evolution. Although we analyzed only a part of the *G* gene, this genotype may have diverged to 4 lineages, including a lineage with new genotype ON2. Genotypes ON1 and ON2 may be potential agents of continuous epidemics of HRSV-A strains in the future.

Acknowledgements

This work was partly supported by a Grant-in-Aid (H25-Shiko-Ippan-015) for Research on Emerging and Re-emerging Infectious Diseases, Labour and Welfare Programs, from the Ministry of Health, Labour and Welfare of Japan.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.meegid.2014.09.030.

References

Agoti, C.N., Otieno, J.R., Gitahi, C.W., Cane, P.A., Nokes, D.J., 2014. Rapid spread and diversification of respiratory syncytial virus genotype ON1, Kenya. Emerg. Infect. 20, 950–959.

Auksornkitti, V., Kamprasert, N., Thongkomplew, S., Suwannakarn, K., Theamboonlers, A., Samransamruajkij, R., Poovorawan, Y., 2013. Molecular characterization of human respiratory syncytial virus, 2010–2011: identification of genotype ON1 and a new subgroup B genotype in Thailand. Arch. Virol. 159, 499–507.

Botosso, V.F., Zanotto, P.M., Ueda, M., Arruda, E., Gilio, A.E., Vieira, S.E., Stewien, K.E., Peret, T.C., Jamal, L.F., Pardini, M.I., Pinho, J.R., Massad, E., Sant'anna, O.A., Holmes, E.C., Durigon, E.L., VGDN Consortium, 2009. Positive selection results in frequent reversible amino acid replacements in the G protein gene of human respiratory syncytial virus. PLoS Pathog. 5, e1000254.

Collins, P.L., Crowe Jr., J.E., 2006. Respiratory syncytial virus and metapneumovirus. In: Knipe, D.M., Howley, P.M., Griffin, D.E., Martin, M.A., Lamb, R.A., Roizman, B., Straus, S.E. (Eds.), Fields Virology, fifth ed. Lippincott Williams & Wilkins, Philadelphia, pp. 1601–1646.

Collins, P.L., Melero, J.A., 2011. Progress in understanding and controlling respiratory syncytial virus: still crazy after all these years. Virus Res. 162, 80–99.

Cui, G., Zhu, R., Qian, Y., Deng, J., Zhao, L., Sun, Y., Wang, F., 2013. Genetic variation in attachment glycoprotein genes of human respiratory syncytial virus subgroups A and B in children in recent five consecutive years. PLoS One 8, e75020.

Dapat, I.C., Shobugawa, Y., Sano, Y., Saito, R., Sasaki, A., Suzuki, Y., Kumaki, A., Zaraket, H., Dapat, C., Oguma, T., Yamaguchi, M., Suzuki, H., 2010. New genotypes within respiratory syncytial virus group B genotype BA in Niigata, Japan. J. Clin. Microbiol. 48, 3423–3427.

de-Paris, F., Beck, C., de Souza, N.L., Machado, A.B., Paiva, R.M., da Silva, M.D., Pires, M.R., dos Santos, R.P., de Souza, K.R., Barth, A.L., 2014. Evaluation of respiratory syncytial virus group A and B genotypes among nosocomial and community-acquired pediatric infections in Southern Brazil. Virol. J. 11, 36.

Domingo, E., Diez, J., Martínez, M.A., Hernandez, J., Holguin, A., Borrego, B., Mateu, M.G., 1993. New observations on antigenic diversification of RNA viruses: antigenic variation is not dependent on immune selection. J. Gen. Virol. 74, 2039–2045.

Domingo, E., 2006. Virus evolution. In: Knipe, D.M., Howley, P.M., Griffin, D.E., Martin, M.A., Lamb, R.A., Roizman, B., Straus, S.E. (Eds.), Fields Virology, fifth ed. Lippincott Willams & Wilkins, Philadelphia, pp. 389–421.

Drummond, A.J., Ho, S.Y., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4, e88.

Drummond, A.J., Rambaut, A., 2007. BEAST: bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

Eshaghi, A., Duvvuri, V.R., Lai, R., Nadarajah, J.T., Li, A., Patel, S.N., Low, D.E., Gubbay, J.B., 2012. Genetic variability of human respiratory syncytial virus A strains circulating in ontario: a novel genotype with a 72 nucleotide *G* gene duplication. PLoS One 7, e32807.

Etemadi, M.R., Sekawi, Z., Othman, N., Lye, M.S., Moghaddam, F.Y., 2013. Circulation of human respiratory syncytial virus strains among hospitalized children with acute lower respiratory infection in Malaysia. Evol. Bioinform. 9, 151–161.

Forcic, D., Ivancic-Jelecki, J., Mlinaric-Galinovic, G., Vojnovic, G., Babic-Erceg, A., Tabain, I., 2012. A study of the genetic variability of human respiratory syncytial virus in Croatia, 2006–2008. J. Med. Virol. 84, 1985–1992.

Gago, S., Elena, S.F., Flores, R., Sanjuán, R., 2009. Extremely high mutation rate of a hammerhead viroid. Science 323, 1308.

Galiano, M.C., Palomo, C., Videla, C.M., Arbiza, J., Melero, J.A., Carballal, G., 2005. Genetic and antigenic variability of human respiratory syncytial virus (groups A and B) isolated over seven consecutive seasons in Argentina (1995 to 2001). J. Clin. Microbiol. 43, 2266–2273.

Hall, C.B., Walsh, E.E., Long, C.E., Schnabel, K.C., 1991. Immunity to and frequency of reinfection with respiratory syncytial virus. J. Infect. Dis. 163, 693–698.

Holmes, E.C., 2011. What does virus evolution tell us about virus origins? J. Virol. 85, 5247–5251.

Kanda, Y., 2013. Investigation of the freely available easy-to-use software 'EZR' for medical statistics. Bone Marrow Transplant. 48, 452–458.

Khor, C.S., Sam, I.C., Hooi, P.S., Chan, Y.F., 2013. Displacement of predominant respiratory syncytial virus genotypes in Malaysia between 1989 and 2011. Infect. Genet. Evol. 14, 357–360.

Kim, Y.J., Kim, D.W., Lee, W.J., Yun, M.R., Lee, H.Y., Lee, H.S., Jung, H.D., Kim, K., 2014. Rapid replacement of human respiratory syncytial virus A with the ON1 genotype having 72 nucleotide duplication in G gene. Infect. Genet. Evol. 26C, 103–112.

Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.

Kushibuchi, I., Kobayashi, M., Kusaka, T., Tsukagoshi, H., Ryo, A., Yoshida, A., Ishii, H., Saraya, T., Kurai, D., Yamamoto, N., Kanou, K., Saitoh, M., Noda, M., Kuroda, M., Morita, Y., Kozawa, K., Oishi, K., Tashiro, M., Kimura, H., 2013. Molecular evolution of attachment glycoprotein (G) gene in human respiratory syncytial virus detected in Japan 2008–2011. Infect. Genet. Evol. 18, 168–173.

Lee, N., Lui, G.C., Wong, K.T., Li, T.C., Tse, E.C., Chan, J.Y., Yu, J., Wong, S.S., Choi, K.W., Wong, R.Y., Ngai, K.L., Hui, D.S., Chan, P.K., 2013. High morbidity and mortality in adults hospitalized for respiratory syncytial virus infections. Clin. Infect. Dis. 57, 1069–1077.

Leung, A.K., Kellner, J.D., Davies, H.D., 2005. Respiratory syncytial virus bronchiolitis. J. Natl Med. Assoc. 97, 1708–1713.

Melero, J.A., García-Barreno, B., Martínez, I., Pringle, C.R., Cane, P.A., 1997. Antigenic structure, evolution and immunobiology of human respiratory syncytial virus attachment (G) protein. J. Gen. Virol. 78, 2411–2418.

Mufson, M.A., Orvell, C., Rafnar, B., Norrby, E., 1985. Two distinct subtypes of human respiratory syncytial virus. J. Gen. Virol. 66, 2111–2124.

Palomo, C., García-Barreno, B., Peñas, C., Melero, J.A., 1991. The G protein of human respiratory syncytial virus: significance of carbohydrate side-chains and the C-terminal end to its antigenicity. J. Gen. Virol. 72, 669–675.

Peret, T.C., Hall, C.B., Schnabel, K.C., Golub, J.A., Anderson, L.J., 1998. Circulation patterns of genetically distinct group A and B strains of human respiratory syncytial virus in a community. J. Gen. Virol. 79, 2221–2229.

Pierangeli, A., Trotta, D., Scagnolari, C., Ferreri, M., Nicolai, A., Midulla, F., Marinelli, K., Antonelli, G., Bagnarelli, P., 2014. Rapid spread of the novel respiratory syncytial virus A ON1 genotype, central Italy, 2011 to 2013. Euro. Surveill. 19, 20843.

Pond, S.L., Frost, S.D., 2005a. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21, 2531–2533.

Pond, S.L., Frost, S.D., 2005b. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol. Biol. Evol. 22, 1208–1222.

Pretorius, M.A., van Niekerk, S., Tempia, S., Moyes, J., Cohen, C., Madhi, S.A., Venter, M., SARI Surveillance Group, 2013. Replacement and positive evolution of

subtype A and B respiratory syncytial virus G-protein genotypes from 1997–2012 in South Africa. J. Infect. 208, S227–237.
Prifert, C., Streng, A., Krempl, C.D., Liese, J., Weissbrich, B., 2013. Novel respiratory syncytial virus A genotype, Germany, 2011–2012. Emerg. Infect. Dis. 19, 1029–1030.
Rebuffo-Scheer, C., Bose, M., He, J., Khaja, S., Ulatowski, M., Beck, E.T., Fan, J., Kumar, S., Nelson, M.I., Henrickson, K.J., 2011. Whole genome sequencing and evolutionary analysis of human respiratory syncytial virus A and B from Milwaukee, WI 1998–2010. PLoS One 6, e25468.
Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.
Sanjuán, R., Nebot, M.R., Chirico, N., Mansky, L.M., Belshaw, R., 2010. Viral mutation rates. J. Virol. 84, 9733–9748.
Shay, D.K., Holman, R.C., Newman, R.D., Liu, L.L., Stout, J.W., Anderson, L.J., 1999. Bronchiolitis-associated hospitalizations among US children, 1980–1996. JAMA 282, 1440–1446.
Shobugawa, Y., Saito, R., Sano, Y., Zaraket, H., Suzuki, Y., Kumaki, A., Dapat, I., Oguma, T., Yamaguchi, M., Suzuki, H., 2009. Emerging genotypes of human respiratory syncytial virus subgroup A among patients in Japan. J. Clin. Microbiol. 47, 2475–2482.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.
Tanabe, A.S., 2011. Kakusan4 and Aminosan: two programs for comparing nonpartitioned, proportional and separate models for combined molecular phylogenetic analyses of multilocus sequence data. Mol. Ecol. Resour. 11, 914–921.
Taubenberger, J.K., Kash, J.C., 2010. Influenza virus evolution, host adaptation, and pandemic formation. Cell Host Microbe 7, 440–451.
Tran, D.N., Pham, T.M., Ha, M.T., Tran, T.T., Dang, T.K., Yoshida, L.M., Okitsu, S., Hayakawa, S., Mizuguchi, M., Ushijima, H., 2013. Molecular epidemiology and disease severity of human respiratory syncytial virus in Vietnam. PLoS One 8, e45436.
Trento, A., Casas, I., Calderón, A., Garcia-Garcia, M.L., Calvo, C., Perez-Breña, P., Melero, J.A., 2010. Ten years of global evolution of the human respiratory syncytial virus BA genotype with a 60-nucleotide duplication in the G protein gene. J. Virol. 84, 7500–7512.
Trento, A., Galiano, M., Videla, C., Carballal, G., García-Barreno, B., Melero, J.A., Palomo, C., 2003. Major changes in the G protein of human respiratory syncytial virus isolates introduced by a duplication of 60 nucleotides. J. Gen. Virol. 84, 3115–3120.
Trento, A., Viegas, M., Galiano, M., Videla, C., Carballal, G., Mistchenko, A.S., Melero, J.A., 2006. Natural history of human respiratory syncytial virus inferred from phylogenetic analysis of the attachment (G) glycoprotein with a 60-nucleotide duplication. J. Virol. 80, 975–984.
Tsukagoshi, H., Yokoi, H., Kobayashi, M., Kushibuchi, I., Okamoto-Nakagawa, R., Yoshida, A., Morita, Y., Noda, M., Yamamoto, N., Sugai, K., Oishi, K., Kozawa, K., Kuroda, M., Shirabe, K., Kimura, H., 2013. Genetic analysis of attachment glycoprotein (G) gene in new genotype ON1 of human respiratory syncytial virus detected in Japan. Microbiol. Immunol. 57, 655–659.
Valley-Omar, Z., Muloiwa, R., Hu, N.C., Eley, B., Hsiao, N.Y., 2013. Novel respiratory syncytial virus subtype ON1 among children, Cape Town, South Africa, 2012. Emerg. Infect. Dis. 19, 668–670.
Venter, M., Madhi, S.A., Tiemessen, C.T., Schoub, B.D., 2001. Genetic diversity and molecular epidemiology of respiratory syncytial virus over four consecutive seasons in South Africa: identification of new subgroup A and B genotypes. J. Gen. Virol. 82, 2117–2124.
Weiss, R.A., McMichael, A.J., 2004. Social and environmental risk factors in the emergence of infectious diseases. Nat. Med. 10, S70–76.
Yamaguchi, M., Sano, Y., Dapat, I.C., Saito, R., Suzuki, Y., Kumaki, A., Shobugawa, Y., Dapat, C., Uchiyama, M., Suzuki, H., 2011. High frequency of repeated infections due to emerging genotypes of human respiratory syncytial viruses among children during eight successive epidemic seasons in Japan. J. Clin. Microbiol. 49, 1034–1040.
Yorita, K.L., Holman, R.C., Steiner, C.A., Effler, P.V., Miyamura, J., Forbes, S., Anderson, L.J., Balaraman, V., 2007. Severe bronchiolitis and respiratory syncytial virus among young children in Hawaii. Pediatr. Infect. Dis. J. 26, 1081–1088.

Jpn. J. Infect. Dis., 67, 326-328, 2014

Laboratory and Epidemiology Communications

An Outbreak of Acute Respiratory Infections due to Human Respiratory Syncytial Virus in a Nursing Home for the Elderly in Ibaraki, Japan, 2014

Ikuko Doi[1], Noriko Nagata[1], Hiroyuki Tsukagoshi[2], Harumi Komori[1], Takumi Motoya[1], Miki Watanabe[1], Toshimasa Keta[1], Manami Kawakami[3], Takashi Tsukano[3], Megumi Honda[3], Taisei Ishioka[4], Makoto Takeda[5], Akihide Ryo[6], Makoto Kuroda[7], Kazunori Oishi[4], and Hirokazu Kimura[4,6]*

[1]*Ibaraki Prefectural Institute of Public Health, Ibaraki 310-0852;*
[2]*Gunma Prefectural Institute of Public Health and Environmental Sciences, Gunma 371-0052;*
[3]*Ibaraki Joso Health Center, Ibaraki 303-0005;*
[4]*Infectious Disease Surveillance Center,* [5]*Department of Virology III, and*
[7]*Pathogen Genomics Center, National Institute of Infectious Diseases, Tokyo 208-0011; and*
[6]*Department of Molecular Biodefence Research, Yokohama City University Graduate School of Medicine, Kanagawa 236-0004, Japan*

Communicated by Masayuki Saijo

(Accepted May 22, 2014)

Human respiratory syncytial virus (HRSV), a member of the family *Paramyxoviridae* and genus *Pneumovirus*, is a notable viral agent that causes acute respiratory infections (ARI) in humans (1). HRSV may also cause severe ARI such as pneumonia in infants (1). However, the epidemiology and pathogenicity of HRSV in elderly persons has not been elucidated. We encountered an outbreak of ARI due to HRSV in a nursing home for the elderly in Ibaraki, Japan during the winter of 2014. Here we report the molecular epidemiological analysis of the outbreak.

Epidemiological investigation suggested that 3 of the 99 residents showed symptoms, such as cough, sore throat, and acute wheezing in the middle of January 2014. They were also diagnosed with pneumonia by chest radiography. Within 9 days, 21 other residents presented with similar symptoms. During this outbreak, the prevalence of infection in the residents was approximately 24% (24/99), but the infection route could not be determined. Patients were aged from 68 to 97 years (81.5 ± 8.5 years; mean ± standard deviation [SD]). Clinical manifestations among the patients were as follows: fever (20/24 residents, 83.3%; 37.7 ± 0.8°C, mean ± SD), rhinorrhea (8/24, 33.3%), cough (21/24, 87.5%), sore throat (7/24, 29.1%), and wheezing (7/24, 29.1%). In total, 5 cases (20.8%) were diagnosed with pneumonia by chest radiography. No underlying conditions including cancer and/or immunosuppressive diseases were observed in all patients with pneumonia. The majority of the patients (21/24) resided on the second floor of the nursing home.

We collected 10 nasopharyngeal swab samples after obtaining verbal informed consent from the patients. We tried to detect and isolate HRSV, influenza A, B, and C viruses, human rhinovirus, enteroviruses, parainfluenza viruses (types 1–4), coronavirus, adenoviruses, *Chlamydophila pneumoniae*, *Mycoplasma pneumoniae*, *Streptococcus pneumoniae*, and *Haemophilus influenzae* using polymerase chain reaction (PCR), reverse transcription (RT)-PCR, or culture methods (2–4). Although HRSV was not isolated using cell culture methods, it was detected from samples by RT-PCR. No other viruses or bacteria were detected or iso-

*Corresponding author: Mailing address: Infectious Disease Surveillance Center, National Institute of Infectious Diseases, 4-7-1 Gakuen, Musashimurayama-shi, Tokyo 208-0011, Japan. Tel: +81-42-561-0771, Fax: +81-42-565-3315, E-mail: kimhiro@nih.go.jp

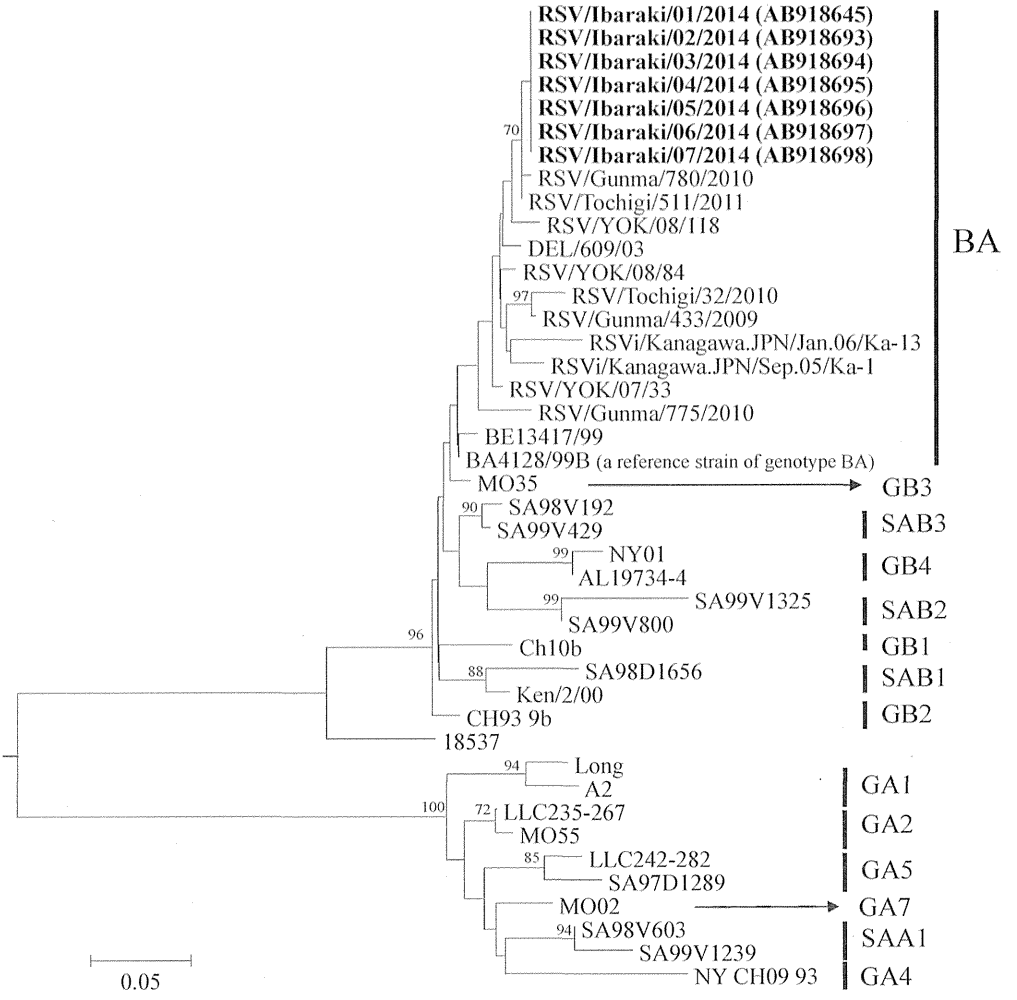

Fig. 1. Phylogenetic tree constructed on the basis of partial sequences of the HRSV *G* gene. Distance was calculated using Kimura's two-parameter method and the tree was plotted using the neighbor-joining method. Numbers above the branches represent the bootstrap probabilities (%). The present strains are shown in bold letters. Numbers in parentheses indicate GenBank accession numbers. The reference strains were as follows: Long (AY911262), A2 (M11486), MO02 (AF233910), NY CH09 93 (AF065254), SA97D1289 (AF348803), SA98V603 (AF348807), SA99V1239 (AF348808), LLC242-282 (AY114150), LLC235-267 (AY114149), MO55 (AF233915), DEL/609/03 (DQ248941), 18537 (M17213), AL19734-4 (AF233924), BA4128/99B (AY333364), BE13417/99 (AY751131), Ch10b (AF065250), CH93 9b (AF065251), Ken/2/00 (AY524575), MO35 (AF233929), NY01 (AF233931), SA98D1656 (AF348826), SA98V192 (AF348811), SA99V429 (AF348813), SA99V800 (AF348821), SA99V1325 (AF348822), RSV/Tochigi/32/2010 (AB775986), RSV/Tochigi/511/2011 (AB775999), RSV/Gunma/433/2009 (AB683222), RSV/Gunma/775/2010 (AB683226), RSV/Gunma/780/2010 (AB683229), RSV/YOK/07/33 (AB551084), RSV/YOK/08/118 (AB551108), RSVi/Kanagawa.JPN/Jan.06/Ka-13 (AB500656), RSVi/Kanagawa.JPN/Sep.05/Ka-1 (AB500659), and RSV/YOK/08/84 (AB551106).

lated. Nucleic acids were extracted from the samples using the QIAamp MinElute Virus Spin Kit (Qiagen, Valencia, CA, USA) and suspended in DNase/RNase-free water. After DNA/RNA extraction, PCR or RT-PCR was performed as described previously (2–4). Amplicons were purified using the QIAquick PCR Purification Kit (Qiagen), and the nucleotide sequences were determined by direct sequencing (5). Next, we performed phylogenetic analysis on the basis of *G* gene nucleotide sequences of HRSV (nucleotide positions: 670–969, 300 nucleotides for the genotype BA reference strain [BA4128B/99B]) using Molecular Evolutionary Genetics Analysis software version 4 (2). Evolutionary distances were estimated using Kimura's two-parameter model, and a phylogenetic tree was constructed using the neighbor-joining method (6,7). The reliability of the tree was estimated using 1000 bootstrap replications. The GenBank accession numbers of the nucleotide sequences obtained are AB918645 and AB918693 to AB918698.

HRSV alone was detected in 7 of the 10 samples collected, and no other pathogens were detected. The nucleotide identity of the analyzed regions (*G* gene) among the present strains was 100%. Phylogenetic analysis based on the HRSV *G* gene nucleotide sequences showed that the strains were HRSV subgroup B (HRSV-B) genotype BA (Fig. 1). In addition, the present strains genetically resembled other domestic HRSV-B genotype BA strains detected in nearby areas (within a 100-km radius) including in Gunma, Tochigi, and Kanagawa prefectures (93.2–99.9% nucleotide identity). All patients recovered without sequelae. Moreover, we carefully examined amino acid substitutions in the C-terminal hypervariable region among the present strains and

other domestic strains. As a result, 2 amino acid substitutions (T259I and T281A) were found. These substitutions might be unique, although further studies are warranted (2,5).

Primary HRSV infection mainly occurs in infants (1). Moreover, HRSV reinfections in the elderly may be associated with severe respiratory infection (pneumonia) or exacerbation of asthma and chronic obstructive pulmonary disease (8). However, the epidemiology of HRSV infection in adults including elderly people is not exactly known. In the present cases, HRSV was detected in 70% of the collected samples. In addition, phylogenetic analysis (Fig. 1) showed that the genotypes of the strains were identical (HRSV-B genotype BA), and the analyzed *G* gene nucleotide sequences completely matched each other. The present strains may be prevalent domestic strains and thus may be closely related genetically. Acute wheezing was observed in approximately 29% (7/24) of patients, and pneumonia was identified in approximately 21% (5/24) of patients. Among them, 4 of the 7 patients presented with pneumonia plus acute wheezing. All 4 of these patients were women, and no chronic pulmonary diseases such as asthma or chronic obstructive pulmonary disease were found. A previous report suggested that wheezing might occur in 6–35% of elderly patients with HRSV infection (8). Thus, constant acute wheezing as a complication of HRSV infection could be observed in the elderly as well as in infants with primary infection of the virus (8–10). In conclusion, HRSV should be considered a possible cause of outbreaks among elderly persons with ARI presenting with pneumonia and acute wheezing.

This article appeared in the Infectious Agents Surveillance Report (IASR), vol. 35, p. 146–147, 2014 in Japanese.

Acknowledgments We thank the staff of the nursing home for their cooperation during the epidemiological investigation.

This work was supported in part by a Grant-in-Aid for Research on Emerging and Re-emerging Infectious Diseases (H25-Shinko-Ippan-015), Labour and Welfare Programs from the Ministry of Health, Labour and Welfare of Japan.

Conflict of interest None to declare.

REFERENCES

1. Collins PL and Karron RA. Respiratory syncytial virus and metapneumovirus. In: Knipe DM, Howley PM, editors. Fields Virology. Vol. 1. 6th ed. Philadelphia: Lippincott Williams & Wilkins; 2013. p. 1086-123.
2. Miyaji Y, Kobayashi M, Sugai K, et al. Severity of respiratory signs and symptoms and virus profiles in Japanese children with acute respiratory illness. Microbiol Immunol. 2013;57:811-21.
3. Abdeldaim GM, Strålin K, Korsgaard J, et al. (2010): Multiplex quantitative PCR for detection of lower respiratory tract infection and meningitis caused by *Streptococcus pneumoniae*, *Haemophilus influenzae* and *Neisseria meningitidis*. BMC Microbiol. 2010;10:310.
4. Thurman KA, Warner AK, Cowart KC, et al. Detection of *Mycoplasma pneumoniae*, *Chlamydia pneumoniae*, and *Legionella* spp. in clinical specimens using a single-tube multiplex real-time PCR assay. Diagn Microbiol Infect Dis. 2011;70:1-9.
5. Yoshida A, Kiyota N, Kobayashi M, et al. Molecular epidemiology of the attachment glycoprotein (G) gene in respiratory syncytial virus in children with acute respiratory infection in Japan in 2009/2010. J Med Microbiol. 2012;61:820-9.
6. Kimura M. A simple method for estimating evolutionary rates of base substitution through comparative studies of nucleotide sequences. J Mol Evol. 1980;16:111-20.
7. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987;4: 406-25.
8. Falsey AR, Walsh EE. Respiratory syncytial virus infection in adults. Clin Microbiol Rev. 2000;13:371-84.
9. Openshaw PJ, Tregoning JS. Immune responses and disease enhancement during respiratory syncytial virus infection. Clin Microbiol Rev. 2005;18:541-55.
10. Silvestri M, Sabatini F, Defilippi AC. The wheezy infant—immunological and molecular considerations. Paediatr Respir Rev. 2004;Suppl A:S81-7.

Int J Clin Exp Pathol 2014;7(8):5308-5312
www.ijcep.com /ISSN:1936-2625/IJCEP0001211

Case Report

Detection of trichodysplasia spinulosa-associated polyomavirus in a fatal case of myocarditis in a seven-month-old girl

Shinya Tsuzuki[1], Hitomi Fukumoto[2], Sohtaro Mine[2,3], Noriko Sato[1], Makoto Mochizuki[3], Hideki Hasegawa[2], Tsuyoshi Sekizuka[4], Makoto Kuroda[4], Takeji Matsushita[1], Harutaka Katano[2]

[1]Department of Pediatrics, National Center for Global Health and Medicine, Tokyo, Japan; [2]Department of Pathology, National Institute of Infectious Diseases, Tokyo, Japan; [3]Department of Pathology, Hospital, National Center for Global Health and Medicine, Tokyo, Japan; [4]Pathogen Genomics Center, National Institute of Infectious Diseases, Tokyo, Japan

Received June 24, 2014; Accepted August 2, 2014; Epub July 15, 2014; Published August 1, 2014

Abstract: Trichodysplasia spinulosa-associated polyomavirus (TSV) was identified in a seven-month-old girl with myocarditis. The number of TSV genomes detected was higher in the heart than in the other organs. The full-length TSV genome was cloned from the heart. This suggests a possible role of TSV infection in the pathogenesis of myocarditis in infants.

Keywords: Trichodysplasia spinulosa-associated polyomavirus, myocarditis

Introduction

Myocarditis is a rare but important cause of sudden death in childhood. Its etiology is wide ranging and often difficult to identify. Viral infection is thought to be one of the most frequent etiological agents of myocarditis [1]. Various viruses, including influenza virus, coxsackievirus B, parvovirus B19, adenoviruses, and cytomegalovirus, have been detected in myocarditis tissue samples. However, it remains unclear whether such viral infections are associated with the pathogenesis of myocarditis in all patients. In this report, we describe an autopsy case of fulminant myocarditis in which trichodysplasia spinulosa-associated polyomavirus (TSV), a human polyomavirus identified in 2010, was detected with next-generation sequencing [2].

Case description

In autumn 2012, a seven-month-old female infant suffering from respiratory discomfort at night attended our emergency outpatient department. She was admitted with acute, severe respiratory distress. She and her family had no appreciable past history, including asthma, congenital heart diseases, and immunodeficiencies. At the initial visit, her bodyweight was 6.3 kg, which was lower than the normal limit of Japanese seven-month-old female infants. Her developmental history was normal. She was appropriately immunized. She did not have any remarkable history of sick contact. Her body temperature was 36.0°C, pulse rate was 132/min, respiratory rate was 70/min, and oxygen saturation was 89% (ambient air). Her blood pressure was not measured. Marked wheezes in both lungs but no abnormal cardiac murmurs were heard on chest auscultation. Chest retraction and skin cyanosis were obvious. Laboratory findings were unremarkable, except for an elevated white blood cell count (14,510/μl) and a reduced hemoglobin level (9.3 g/dl). A venous blood gas analysis demonstrated mild acidemia and hypercapnea (pH 7.252, pCO_2 45.6 mmHg, base excess -6.9 mEq/l, HCO_3^- 19.4 mEq/l). No immunological test could be conducted because the patient came to the hospital at night, and no blood sample was preserved due to difficulty drawing

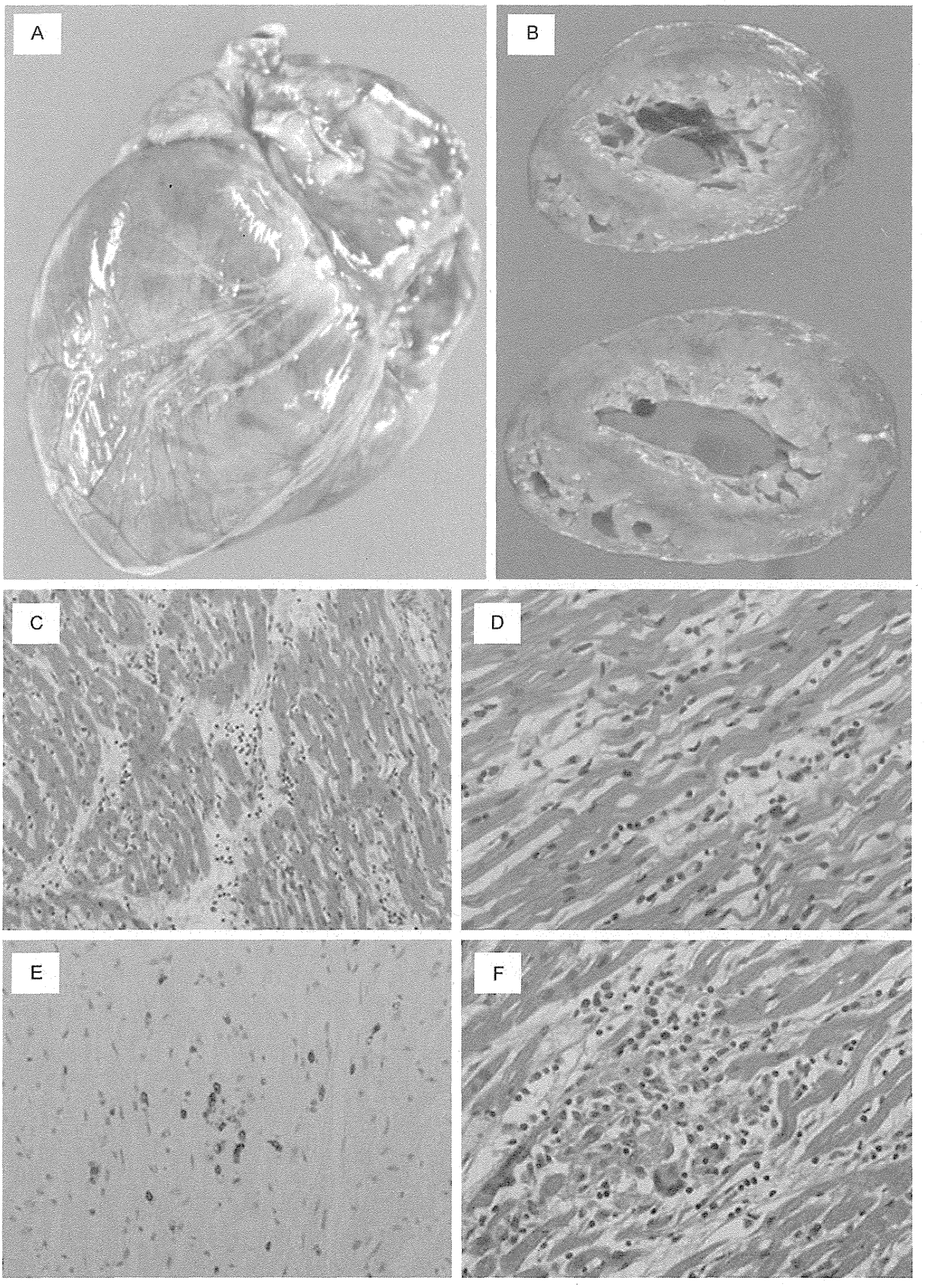

Figure 1. Macroscopic and histological findings in the heart. A, B. Macroscopic view of the heart. A. Petechial hemorrhage was observed on the surface. B. Left ventricle was enlarged and small petechiae were found on the cut surface. C-F. Histological views of the heart. C, D. Hematoxylin and eosin staining of the myocardial tissue. The

myocardial bundles were disturbed and edematous. Edema with infiltration by small lymphocytes and macrophages was observed between the myocardial bundles. E. Immunohistochemical analysis of CD8. F. Small granulomas with giant cells were found in the myocardium.

blood. Blood culture was conducted at the time of autopsy and demonstrated no significant bacterial growth. A chest radiograph showed abnormal infiltration in the right hilar area and lower lung field, and an increased cardiothoracic ratio. The lateral view also showed a posteriorly displaced major bronchus.

The patient was administered oxygen and inhaled salbutamol in the emergency room. Her respiratory rate and mode of respiration improved after inhalation, and she was diagnosed with a severe asthmatic attack complicated with a suspected lower respiratory infection and admitted to a pediatric ward. Three hours after admission, her respiratory condition deteriorated. Transfer to the intensive-care unit was determined and additional examinations and intubation was performed. A blood test revealed an extremely high brain natriuretic peptide level (6,788 pg/ml), whereas her creatine phosphokinase, lactate dehydrogenase, and aspartate aminotransferase levels were in the normal ranges. An echocardiogram was unremarkable, but an electrocardiogram was not performed. Cardiac arrest occurred suddenly after intubation. Resuscitation efforts were unsuccessful and the patient died 8 hour after admission. A morbid autopsy was performed within 6 hours of death. Macroscopically, the patient's heart was large and the left ventricle was dilated. There was no thrombus in the lumen or significant change in the four valves. Severe diffuse infiltration of small lymphocytes and macrophages was observed histologically in the myocardium (**Figure 1**). Immunohistochemistry showed that a large proportion of the small lymphocytes were $CD8^+$ T cells. Small granulomas with giant cells were found in the myocardium. There was no necrosis, and no disturbance of the myocardial array was observed. Small granulomas were also found in the lung and liver. Neither fungi nor acid-fast bacilli were found. In the both lungs, diffuse and mild lymphoplasmacytic infiltration of the bronchial and alveolar wall was observed.

Materials and methods

To identify any pathogen causing myocarditis, DNA and RNA samples extracted from a frozen sample of the heart were analyzed with a multi-virus real-time PCR system, which can examine 163 viruses simultaneously on a 96-well plate [3]. Although a low copy number of parvovirus B19 (36.1 copies per 100 ng of DNA) was detected in the bronchus, no viral genome was detected in the heart sample with this system. We then analyzed DNA and RNA samples extracted from frozen tissue from the heart with deep sequencing, using a next-generation sequencer (GAIIx, Illumina, San Diego, CA). RNA was reverse transcribed to cDNA, and both DNA and cDNA were applied to the GAIIx. In total, 3×10^7 reads were obtained with the GAIIx, and a BLAST search revealed that 244,965 80-bp reads were nonhuman sequences.

Results

Two of the reads corresponded to nucleotides (nt) 4,535-4,614 and nt 5,034-5,113 of the TSV genome (GU989205). No other significant pathogenic genome was identified. A real-time PCR analysis was performed to detect the TSV VP1 gene using the TSPyV-F primer (5'-cagtgctaatgacaaattggttgttc-3'), TSPyV-R primer (5'-ttagcttttgtttgtagtgaggattga-3'), and TSPyV-FAM probe (5'-FAM-cccaataaaacaccagagcacacaaggc-TAMRA-3'), targeting nt 1,841-1,923 of GU989205. The real-time PCR analysis detected the TSV genome in the heart (413.2 copies per 100 ng of DNA), lung (248.8), liver (81.6), spleen (103.7), bronchus (36.5), small intestine (58.1), and colon (24.1), indicating that the TSV genome copy number was higher in the heart than in the other organs.

Long PCR using KOD-FX DNA polymerase (Toyobo, Tokyo, Japan) amplified the entire TSV genome, with a length of 5,232 bp. We also successfully PCR amplified a 252-bp fragment including a putative replication origin of TSV using two primers (TSV-F5111 5'-agcctctgtgtgcctcaatt-3', and TSV-R131 5'-tcctcaggataacggtcttaa-3') suggesting the presence of a circular form of the TSV genome in the heart sample. The long PCR product of full-length TSV was cloned into the pCR2.1-blunt vector (Invitrogen, Carlsbad, CA). A sequencing analysis of four clones from independent colonies revealed

that the cloned PCR product was 99.4% homologous to the complete genome of previously reported TSV (GU989205 and JQ723730). Therefore, we designated the TSV clone strain TSV-TMC, and registered it in GenBank under accession number AB873001.

Discussion

Polyomaviruses are ~45 nm, nonenveloped, double-stranded DNA viruses with a small genome of ~5.2 kb. Until now, 10 polyomaviruses have been described in humans and the majority of human polyomaviruses (HPyVs) were identified in the last five years [4]. Generally, a primary infection with polyomavirus is thought to be asymptomatic, and the virus persists in immunocompetent individuals as a latent infection during their whole lives. JC polyomavirus and BK polyomavirus are considered to be pathogens that cause progressive multifocal leukoencephalopathy in AIDS patients and nephropathy in renal transplant recipients, respectively. Merkel cell polyomavirus (MCPyV) has been found to be integrated in a large proportion of Merkel cell carcinomas of the skin [5]. HPyV6 and HPyV7 productively infect the human skin, and infections with such skin-tropic polyomaviruses are very common in the general population [6]. A putative association between trichodysplasia spinulosa (TS) and TSV has been considered, but is not definitive [7]. TS is histologically characterized by the abnormal maturation and marked distention of the hair follicles, and is considered to be symptomatic only in immunocompromised patients. However, the actual etiology of TSV has not yet been established.

TSV was detected in the myocardial tissues of a seven-month-old girl with severe myocarditis. This report is notable for two points: (1) TSV was identified in a child who had no remarkable past history and would be immunocompetent; and (2) TSV was detected in tissues other than the skin. Although the seroprevalence of TSV in the general Japanese population is unknown, a high prevalence is predicted in Japan based on the seroprevalence data from European countries [8-11]. It has also been demonstrated that primary TSV infections frequently occur in early life, within 0-10 years after birth [8]. However, no report has described an association between primary TSV infection and any disease. We were unable to determine whether this was a primary infection, but the high copy number of the TSV genome and the presence of the circular form of TSV in the heart suggest a possible role of TSV as a causative agent of myocarditis in some infant cases. Although the molecular mechanism of TSV infection is unknown, this case serves a hypothesis of the presence of specific receptor for TSV in myocytes. Alternatively, TSV might be reactivated and produced in specific organs under the condition of severe inflammation. Histological changes in this case such as severe infiltration of lymphocytes and destruction of myocytes are observed commonly among other viral myocarditis like Coxsackievirus and enterovirus, not specific for TSV. It cannot be completely denied that such severe inflammation induced production of TSV which existed as a bystander there. To determine the association between myocarditis and TSV infection, serological studies and the immunohistochemical detection of the viral proteins in myocarditis samples will be required. Immunohistochemistry and in situ hybridization may be useful for detection of the virus protein or genome in tissue samples, but have not been established, yet, because there is no appropriate control sample or culture cell positive for TSV, so far. The direct relationship between granulomas and TSV infection was considered unlikely, because distribution of the granuloma was not correlated with TSV copy numbers. However, the presence of granulomas implies any immunological abnormality in the patient.

To the best of our knowledge, this is the first case report describing the detection of TSV in a myocarditis specimen from a child. Although we have previously reported that low copy numbers of MCPyV were detected in some tissue samples from myocarditis patients [12], few reports have described the association between polyomavirus infection and myocarditis. Therefore, further studies are required to clarify the association between TSV infection and the pathogenesis of myocarditis.

Acknowledgements

The authors deeply appreciate the sincere cooperation of the parents of the patient. This work was partly supported by Health and Labor Sciences Research Grants (No. H25-Shinko-Ippan 015 to MK and HK) from the Ministry of Health, Labor and Welfare, and a Grant-in-Aid

for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (No. 24659212 to HK).

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Shinya Tsuzuki, Department of Pediatrics, National Center for Global Health and Medicine, 1-21-1 Toyama, Shinjuku-ku, Tokyo 162-0855, Japan. Tel: +81(3)-3202-7181; Fax: +81(3)-3207-1038; E-mail: shinyatsuzuki@yahoo.co.jp

References

[1] Cooper LT Jr. Myocarditis. N Engl J Med 2009; 360: 1526-1538.

[2] van der Meijden E, Janssens RW, Lauber C, Bouwes Bavinck JN, Gorbalenya AE, Feltkamp MC. Discovery of a new human polyomavirus associated with trichodysplasia spinulosa in an immunocompromized patient. PLoS Pathog 2010; 6: e1001024.

[3] Katano H, Kano M, Nakamura T, Kanno T, Asanuma H, Sata T. A novel real-time PCR system for simultaneous detection of human viruses in clinical samples from patients with uncertain diagnoses. J Med Virol 2011; 83: 322-330.

[4] Kazem S, van der Meijden E, Feltkamp MC. The trichodysplasia spinulosa-associated polyomavirus: virological background and clinical implications. APMIS 2013; 121: 770-782.

[5] Feng H, Shuda M, Chang Y, Moore PS. Clonal integration of a polyomavirus in human Merkel cell carcinoma. Science 2008; 319: 1096-1100.

[6] Schowalter RM, Pastrana DV, Pumphrey KA, Moyer AL, Buck CB. Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin. Cell Host Microbe 2010; 7: 509-515.

[7] Kazem S, van der Meijden E, Kooijman S, Rosenberg AS, Hughey LC, Browning JC, Sadler G, Busam K, Pope E, Benoit T, Fleckman P, de Vries E, Eekhof JA, Feltkamp MC. Trichodysplasia spinulosa is characterized by active polyomavirus infection. J Clin Virol 2012; 53: 225-230.

[8] Nicol JT, Robinot R, Carpentier A, Carandina G, Mazzoni E, Tognon M, Touzé A, Coursaget P. Age-specific seroprevalences of Merkel cell polyomavirus, human polyomaviruses 6, 7, and 9, and trichodysplasia spinulosa-associated polyomavirus. Clin Vaccine Immunol 2013; 20: 363-368.

[9] Sadeghi M, Aronen M, Chen T, Jartti L, Jartti T, Ruuskanen O, Söderlund-Venermo M, Hedman K. Merkel cell polyomavirus and trichodysplasia spinulosa-associated polyomavirus DNAs and antibodies in blood among the elderly. BMC Infect Dis 2012; 12: 383.

[10] Chen T, Mattila PS, Jartti T, Ruuskanen O, Soderlund-Venermo M, Hedman K. Seroepidemiology of the newly found trichodysplasia spinulosa-associated polyomavirus. J Infect Dis 2011; 204: 1523-1526.

[11] van der Meijden E, Kazem S, Burgers MM, Janssens R, Bouwes Bavinck JN, de Melker H, Feltkamp MC. Seroprevalence of trichodysplasia spinulosa-associated polyomavirus. Emerg Infect Dis 2011; 17: 1355-1363.

[12] Fukumoto H, Sato Y, Hasegawa H, Katano H. Frequent detection of Merkel cell polyomavirus DNA in sera of HIV-1-positive patients. Virol J 2013; 10: 84.