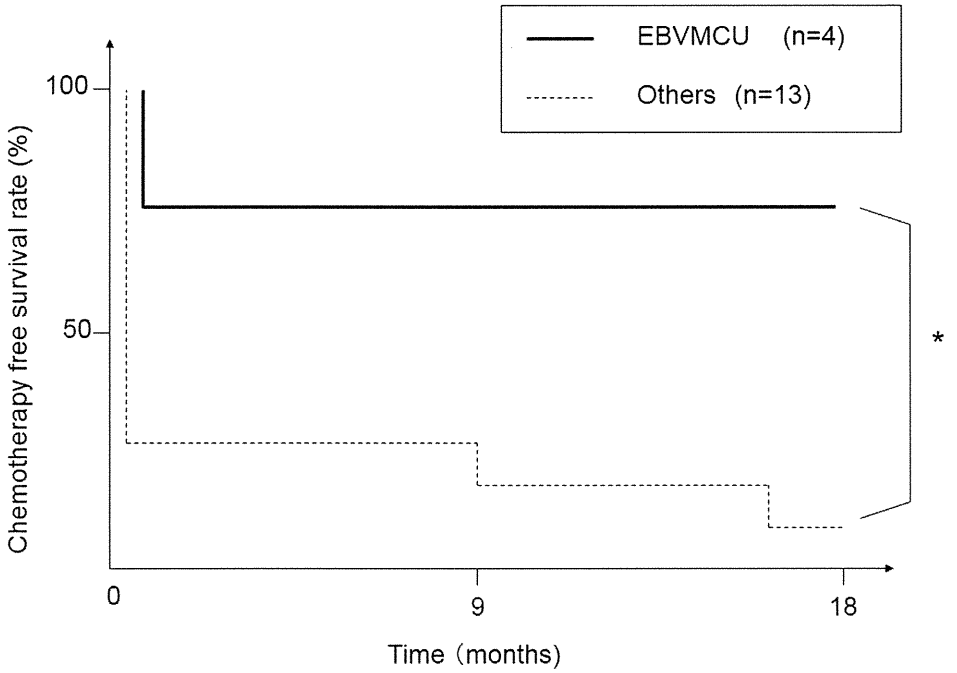

Figure 3. Chemotherapy-free survival. Kaplan-Meier curve shows chemotherapy-free survival within 18 months, comparing patients in the EBV-positive mucocutaneous ulcer (EBVMCU) group with those with other forms of MTX-LPD. Methotrexate was withdrawn in all cases. *Log-rank test, $p < 0.05$. MTX-LPD: methotrexate-associated lymphoproliferative disorders; EBV: Epstein-Barr virus.

Table 4. EBV latency classification among MTX-LPD, PT-LPD, and Age-LPD.

Disease	Year	Author	Cases	I	II	III
PT-LPD	2003	Birkeland	16	1	8	7
Age-LPD	2009	Asano	26	0	19	7
MTX-LPD	2007	Miyazaki	3	0	1	2
MTX-LPD	2013	present case series	11	3	5	3

MTX-LPD: methotrexate-associated lymphoproliferative disorders; PT-LPD: posttransplant lymphoproliferative disorders.

likely to be in latency II, followed by latency III. This observation also appears to hold true for PT-LPD and Age-LPD. PT-LPD has been generally categorized as latency III[25], but is often seen to be in latency II. Cases of Age-LPD were also found to be mainly in latency II. Nonetheless, the latency classification had no value in predicting the prognosis of MTX-LPD in our cohort. The utility of EBV latency in MTX-LPD remains unknown.

Although the pathogenesis of EBVMCU is unclear, lymphocytic vasculitis was observed in more than half our cases of EBVMCU. In EBV-positive LPD, the Mig monokine, induced by interferon-γ (IFN-γ), and IFN-γ inducible protein-10 (mainly produced by reactive cells including endothelial cells) are thought to be powerful instigators of vascular and tissue injuries[26]. Thus, we hypothesize that tissue necrosis and the impairment of local blood flow to the area of vascular damage might be pivotal to the pathogenesis of EBVMCU. Although it is unclear why patients with MTX-LPD who developed EBVMCU had a better prognosis, several factors might be responsible. One possibility is that the mucocutaneous ulcer is so conspicuous that patients seek medical help more promptly, when diagnosis is relatively straightforward and the disease is in an earlier, potentially reversible stage. Thus, the prompt cessation of MTX may lead to a good outcome.

Our study is the first, to our knowledge, to have conducted HLA typing, revealing that HLA-B15:11 could be a risk allele in EBV+ RA with MTX-LPD. Notably, all 3 HLA-B15:11-positive cases were EBV-positive and 2 were polymorphic, indicating that this allele may be linked to the susceptibility to EBV infection and development of LPD. In our study, HLA-B15:11 had no correlation with HLA-shared epitopes, RF, or ACPA. To the best of our knowledge, there have been no reports of any relationship between this allele and other diseases. The identification of a genetic risk factor could help to clarify the pathogenesis of

LPD and to achieve safer therapy for patients who might be at risk of LPD if their RA is treated with MTX. Because this is a retrospective study, there are some limitations, and a larger scale prospective trial is needed to clarify the pathogenesis of this disease.

We have demonstrated that cases of EBVMCU, a subgroup of MTX-LPD, were all histologically polymorphic and had a more favorable outcome by withdrawing MTX alone. In addition, we found that the frequency of the HLA-B15:11 allele was significantly increased in our cohort, which suggests that it may be a risk factor for EBV+ RA with MTX-LPD.

ACKNOWLEDGMENT

The authors thank Dr. Toda (Kyoto University, Center for Anatomical Studies) for performing immunostaining studies and Dr. Inoko (Tokai University, Department of Molecular Science) for kindly providing the HLA reference data of Japanese patients with rheumatoid arthritis.

REFERENCES

1. Gaulard P. Swerdlow SH, Harris NL, Jaffe ES, Sundstrom C. Other iatrogenic immunodeficiency-associated lymphoproliferative disorders. In: Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, et al, eds. Tumours of haematopoietic and lymphoid tissues. Pathology & genetics. World Health Organization classification of tumours. 4th ed. Lyon: IARC Press; 2008:350-1.
2. Weinblatt ME, Coblyn JS, Fox DA, Fraser PA, Holdsworth DE, Glass DN, et al. Efficacy of low-dose methotrexate in rheumatoid arthritis. N Engl J Med 1985;312:818-22.
3. Rath T, Rubbert A. Drug combinations with methotrexate to treat rheumatoid arthritis. Clin Exp Rheumatol 2010;5 Suppl 61:S52-7.
4. Harris NL, Swerdlow SH. Methotrexate-associated lymphoproliferative disorders. In: Jaffe ES, Harris NL, Stein H, Vardiman JW, eds. Tumours of haematopoietic and lymphoid tissues. Pathology & genetics. World Health Organization classification of tumours. 3rd ed. Lyon: IARC Press; 2001:270-1.
5. Feng WH, Cohen JI, Fischer S, Li L, Sneller M, Goldbach-Mansky R, et al. Reactivation of latent Epstein-Barr virus by methotrexate: a potential contributor to methotrexate-associated lymphomas. J Natl Cancer Inst 2004;96:1691-702.
6. Young LS, Rickinson AB. Epstein-Barr virus: 40 years on. Nat Rev Cancer 2004;4:757-68.
7. Dojcinov SD, Venkataraman G, Raffeld M, Pittaluga S, Jaffe ES. EBV positive mucocutaneous ulcer—a study of 26 cases associated with various sources of immunosuppression. Am J Surg Pathol 2010;34:405-17.
8. Oyama T, Ichimura K, Suzuki R, Suzumiya J, Ohshima K, Yatabe Y, et al. Senile EBV+ B-cell lymphoproliferative disorders: a clinicopathologic study of 22 patients. Am J Surg Pathol 2003;27:16-26.
9. Ohmura K, Terao C, Maruya E, Katayama M, Matoba K, Shimada K, et al. Anti-citrullinated peptide antibody-negative RA is a genetically distinct subset: a definitive study using only bone-erosive ACPA-negative rheumatoid arthritis. Rheumatology 2010;49:2298-304.
10. Hashizume H, Uchiyama I, Kawamura T, Suda T, Takigawa M, Tokura Y. Epstein-Barr virus-positive mucocutaneous ulcers as a manifestation of methotrexate-associated B-cell lymphoproliferative disorders. Acta Derm Venereol 2012;92:276-7.
11. Curry JL, Prieto VG, Jones DM, Vega F, Duvic M, Diwan AH. Transient iatrogenic immunodeficiency-related B-cell lymphoproliferative disorder of the skin in a patient with mycosis fungoides/Sezary syndrome. J Cutan Pathol 2011;38:295-7.
12. Shimura C, Satoh T, Takayama K, Yokozeki H. Methotrexate-related lymphoproliferative disorder with extensive vascular involvement in a patient with rheumatoid arthritis. J Am Acad Dermatol 2009;61:126-9.
13. Pastor-Nieto MA, Kilmurray LG, Lopez-Chumillas A, O'Valle F, Garcia-Del Moral R, Puig AM, et al. Methotrexate-associated lymphoproliferative disorder presenting as oral ulcers in a patient with rheumatoid arthritis [in Spanish]. Actas Dermosifiliogr 2009;100:142-6.
14. Uneda S, Sonoki T, Nakamura Y, Matsuoka H, Nakakuma H. Rapid vanishing of tumors by withdrawal of methotrexate in Epstein-Barr virus-related B cell lymphoproliferative disorder. Intern Med 2008;47:1445-6.
15. Tanaka A, Shigematsu H, Kojima M, Sakashita H, Kusama K. Methotrexate-associated lymphoproliferative disorder arising in a patient with adult Still's disease. J Oral Maxillofac Surg 2008;66:1492-5.
16. Maruani A, Wierzbicka E, Machet MC, Abdallah-Lotf M, de Muret A, Machet L. Reversal of multifocal cutaneous lymphoproliferative disease associated with Epstein-Barr virus after withdrawal of methotrexate therapy for rheumatoid arthritis. J Am Acad Dermatol 2007;5 Suppl:S69-71.
17. Clarke LE, Junkins-Hopkins J, Seykora JT, Adler DJ, Elenitsas R. Methotrexate-associated lymphoproliferative disorder in a patient with rheumatoid arthritis presenting in the skin. J Am Acad Dermatol 2007;56:686-90.
18. Dojcinov SD, Venkataraman G, Pittaluga S, Wlodarska I, Schrager JA, Raffeld M, et al. Age-related EBV-associated lymphoproliferative disorders in the Western population: a spectrum of reactive lymphoid hyperplasia and lymphoma. Blood 2011;117:4726-35.
19. Oyama T, Yamamoto K, Asano N, Oshiro A, Suzuki R, Kagami Y, et al. Age-related EBV-associated B-cell lymphoproliferative disorders constitute a distinct clinicopathologic group: a study of 96 patients. Clin Cancer Res 2007;13:5124-32.
20. Shimoyama Y, Yamamoto K, Asano N, Oyama T, Kinoshita T, Nakamura S. Age-related Epstein-Barr virus-associated B-cell lymphoproliferative disorders: special references to lymphomas surrounding this newly recognized clinicopathologic disease. Cancer Sci 2008;99:1085-91.
21. Miyazaki T, Fujimaki K, Shirasugi Y, Yoshiba F, Ohsaka M, Miyazaki K, et al. Remission of lymphoma after withdrawal of methotrexate in rheumatoid arthritis: relationship with type of latent Epstein-Barr virus infection. Am J Hematol 2007;82:1106-9.
22. Birkeland SA, Hamilton-Dutoit S. Is posttransplant lymphoproliferative disorder (PTLD) caused by any specific immunosuppressive drug or by the transplantation per se? Transplantation 2003;76:984-8.
23. Asano N, Yamamoto K, Tamaru J, Oyama T, Ishida F, Ohshima K, et al. Age-related Epstein-Barr virus (EBV)-associated B-cell lymphoproliferative disorders: comparison with EBV-positive classic Hodgkin lymphoma in elderly patients. Blood 2009;113:2629-36.
24. Hoshida Y, Tomita Y, Zhiming D, Yamauchi A, Nakatsuka S, Kurasono Y, et al. Lymphoproliferative disorders in autoimmune diseases in Japan: analysis of clinicopathological features and Epstein-Barr virus infection. Int J Cancer 2004;108:443-9.
25. Masucci MG. Epstein-Barr virus oncogenesis and the ubiquitin-proteasome system. Oncogene 2004;23:2107-15.
26. Teruya-Feldstein J, Jaffe ES, Burd PR, Kanegane H, Kingma DW, Wilson WH, et al. The role of Mig, the monokine induced by interferon-gamma, and IP-10, the interferon-gamma-inducible protein-10, in tissue necrosis and vascular damage associated with Epstein-Barr virus-positive lymphoproliferative disease. Blood 1997;90:4099-105.

DOI:10.1093/jnci/djt338
Advance Access publication December 7, 2013

BRIEF COMMUNICATION

Novel Germline Mutation in the Transmembrane Domain of *HER2* in Familial Lung Adenocarcinomas

Hiromasa Yamamoto, Koichiro Higasa, Masakiyo Sakaguchi, Kazuhiko Shien, Junichi Soh, Koichi Ichimura, Masashi Furukawa, Shinsuke Hashida, Kazunori Tsukuda, Nagio Takigawa, Keitaro Matsuo, Katsuyuki Kiura, Shinichiro Miyoshi, Fumihiko Matsuda, Shinichi Toyooka

Manuscript received July 7, 2013; revised October 14, 2013; accepted October 16, 2013.

Correspondence to: Shinichi Toyooka, MD, PhD, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Clinical Genomic Medicine/Thoracic, Breast and Endocrinological Surgery, 2-5-1 Shikata-cho, Kita-ku, Okayama, Okayama 700–8558, Japan (e-mail: toyooka@md.okayama-u.ac.jp).

We encountered a family of Japanese descent in which multiple members developed lung cancer. Using whole-exome sequencing, we identified a novel germline mutation in the transmembrane domain of the human epidermal growth factor receptor 2 (*HER2*) gene (G660D). A novel somatic mutation (V659E) was also detected in the transmembrane domain of *HER2* in one of 253 sporadic lung adenocarcinomas. Because the transmembrane domain of *HER2* is considered to be responsible for the dimerization and subsequent activation of the *HER* family and downstream signaling pathways, we performed functional analyses of these *HER2* mutants. Mutant HER2 G660D and V659E proteins were more stable than wild-type protein. Both the G660D and V659E mutants activated Akt. In addition, they activated p38, which is thought to promote cell proliferation in lung adenocarcinoma. Our findings strongly suggest that mutations in the transmembrane domain of *HER2* may be oncogenic, causing hereditary and sporadic lung adenocarcinomas.

J Natl Cancer Inst (2014) 106(1): djt338

Familial lung cancers are rare among human malignancies. Recent studies have reported that germline mutations in the epidermal growth factor receptor (*EGFR*) gene predispose the development of lung cancer. Reported familial lung adenocarcinomas with a germline *EGFR* mutation, such as T790M, carry secondary somatic *EGFR* mutations, including exon 19 deletion and exon 21 L858R mutation (1–4). We encountered a family of Japanese descent in which multiple members developed lung cancer (Figure 1). The proband (III-4) was a 53-year-old woman with multiple lung adenocarcinomas in bilateral lungs. She was a light smoker with a 1.2-pack-year history of smoking. She had undergone a left lower lobectomy for multiple lung adenocarcinomas at the age of 44 years. Her mother (II-4), a never smoker, also had multiple lung adenocarcinomas. Partial pulmonary resections of two tumors were performed for II-4 for the purpose of diagnosis after pleural dissemination was found during surgery, and multiple lesions were removed in a lobectomy or partial resections in III-4. A histological examination of the resected tumors in II-4 revealed nonmucinous adenocarcinoma in situ and nonmucinous minimally invasive adenocarcinoma, whereas

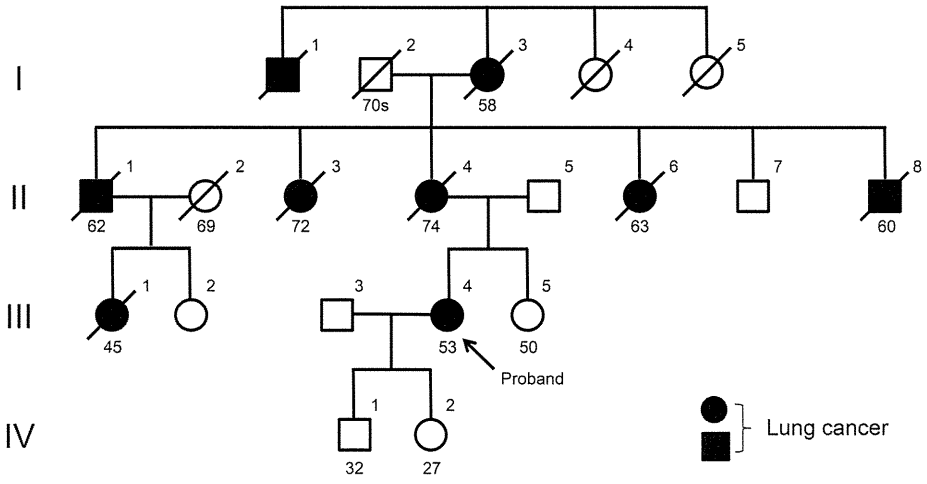

Figure 1. Pedigree chart of a Japanese family in which multiple members developed lung cancer. The **boxes** and **circles** indicate men and women, respectively. The **numbers** at the bottom of each member indicate the age at the time of death or the time of the analysis. An **oblique line** shows deceased family members. The proband (III-4) had multiple lung adenocarcinomas (**arrow**). Tumor tissue, nonmalignant lung tissue, and peripheral blood samples were obtained from III-4. The proband's mother (II-4) also had multiple lung adenocarcinomas, and tumor and nonmalignant lung tissue samples were available. The proband's father (II-5) and sister (III-5) were both unaffected, and peripheral blood samples were obtained from these individuals. Some family members who were not considered as critical for this study were excluded from the pedigree chart to preserve confidentiality. Whole-exome sequencing was performed for individuals II-4, II-5, III-4, and III-5.

the histological findings of pleural dissemination indicated mucus-containing adenocarcinoma. Those of III-4 contained various subtypes of adenocarcinoma, including nonmucinous and mucinous adenocarcinoma in situ and invasive mucinous adenocarcinoma. In addition, normal-appearing lung parenchyma obtained from a lobectomy in III-4 revealed innumerable small preinvasive lesions, implying the presence of precancerous changes throughout the lung (Supplementary Figure 1, available online). Sequencing analyses of *EGFR* exons 18 to 21 and *KRAS* as well as an immunohistochemical staining for ALK protein in the resected tumors indicated no genetic alterations in these genes. The pedigree chart suggested that lung cancer was inherited in an autosomal dominant manner.

After obtaining permission from the Institutional Review Board at Okayama University Hospital and informed consent from the patients and other family members, we performed a whole-exome sequencing study. Tumor DNA samples from II-4, tumor and peripheral blood DNA samples from III-4, and peripheral blood DNA samples from two unaffected family members (II-5 and III-5) were used for the analysis. The candidate germline alterations were restricted to 29 variants by comparing the whole-exome sequencing results between the patients and the unaffected family members. Among them, we focused on a point mutation in the human epidermal growth factor receptor 2 (*HER2/neu*) gene (NM_004448, G660D, GGC to GAC), which was located in exon 17 encoding the transmembrane domain of *HER2* (Supplementary Tables 1–3). This alteration was confirmed by direct sequencing (Figure 2A). We also confirmed that there was no copy number gain of *HER2* in the examined tumors based on the degree of read-depth in the whole-exome sequencing results. Of note, no mutations in genes known to cause lung cancers were detected for tumors from III-4 and II-4.

We considered that somatic mutations in the *HER2* transmembrane domain might act as driver mutations in lung cancer. Hence, we sequenced exon 17 of the *HER2*

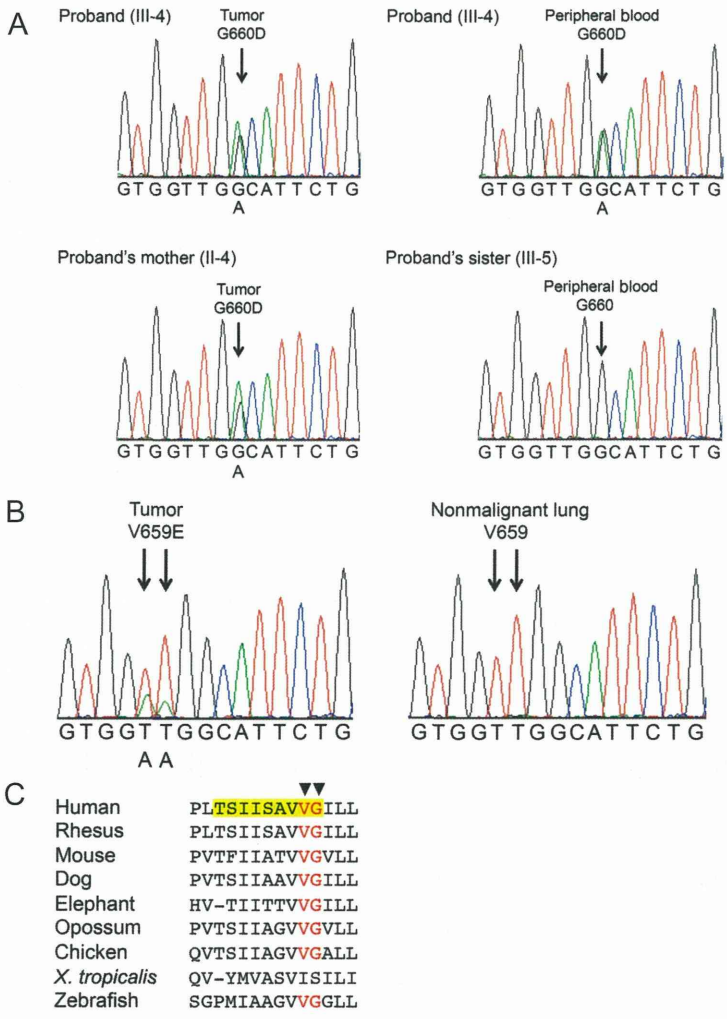

Figure 2. DNA and amino acid sequences in the transmembrane domain of *HER2*. **A**) Direct Sanger sequencing of the proband (III-4), her affected mother (II-4), and her unaffected sister (III-5). The results indicated that G660D was a germline mutation. **B**) Direct sequencing of a sporadic lung adenocarcinoma with a *HER2* V659E mutation. V659E was found to be of somatic origin based on the sequencing results of the peritumoral lung tissue from the same specimen. All the sequence variants were confirmed by independent polymerase chain reaction amplifications and were sequenced in both directions. **C**) Interspecies conservation of the transmembrane domain of *HER2* (UCSC Genome Browser, http://genome.ucsc.edu, accessed September 12, 2013). The **yellow highlight** indicates the N-terminal glycine zipper motif Thr^{652}-X_3-Ser^{656}-X_3-Gly^{660}, a tandem variant of a GG4-like motif of human *HER2*. Codons 659 and 660 in human *HER2* are highly conserved among the listed vertebrate species (shown in **red**). *X. tropicalis* = *Xenopus tropicalis*.

in the tumor samples of 315 sporadic non–small cell lung cancer patients, of which 253 were adenocarcinomas. Although the *HER2* G660D mutation was not detected, a novel nonsynonymous mutation, V659E (GTT to GAA), next to codon 660 was identified in one of these patients. This patient was histologically diagnosed as nonmucinous adenocarcinoma in situ, and the patient had neither smoking history nor apparent family history of lung cancer. This V659E mutation was certainly a somatic mutation because it was not identified in the peritumoral lung tissue of the same patient (Figure 2B). The alignment of *HER2* amino acid sequences showed high conservation of valine 659 and glycine 660 among vertebrates (Figure 2C).

HER2 somatic mutations have been reported in 2% to 4% of lung adenocarcinomas (5–7). However, all reported mutations were restricted to its tyrosine kinase domain (6,7). According to the cBioPortal for Cancer Genomics (http://www.cbioportal.org/public-portal/, accessed September 12, 2013), the same genetic mutation in the *HER2* has not been reported in any type of cancer. Interestingly, a previous study reported that a mutation in the transmembrane domain (V664E) of the rat *neu* gene, which corresponds to V659E in its human homolog *HER2*, induced oncogenic transformation (8). In addition, in vivo experiments showed that the *HER2* V659E mutation contributed to the stability of HER2 dimers, resulting in the dysregulated receptor activation and subsequent cell transformation (9,10). Furthermore, the novel mutations were located within the glycine zipper motif Thr^{652}-X_3-Ser^{656}-X_3-Gly^{660}, a tandem variant of the GG4-like motif, at the N-terminal portion of the transmembrane domain, which was critically related to the dimerization of HER2 (Figure 2C) (9,11). Accordingly, we performed a functional analysis of the mutant HER2 proteins. We found that the degradation of HER2 protein after the administration of cycloheximide was slower in G660D and V659E mutants as compared with wild-type (Supplementary Figure 2A), indicating the higher stability of the mutant proteins than wild-type protein. In addition, results of a phospho-mitogen–activated protein kinase array indicated the activation of Akt and p38α (data not shown). Indeed, Akt is known to be activated by HER2 by phosphatidylinositol 3-kinase and leads to increased cell growth and survival (12,13). Also, the activation of p38 was shown to contribute to the viability of lung adenocarcinoma cells derived from never or light smokers (14,15). A western blot analysis for Akt and p38 successfully confirmed the upregulation of both phospho-Akt and phospho-p38 expression in the mutant *HER2* transfectants (Supplementary Figure 2B).

Because the G660D alteration in *HER2* might have been the cause of the lung cancer in the pedigree studied, we investigated whether familial aggregation of cancer in other organs could be seen in this pedigree. We found that II-1 and II-6 developed renal and gastric cancers, respectively; however, both of them also had lung cancer. The reason why other types of clinically apparent malignances were rarely found in this pedigree is unclear. The G660D germline mutation may be tolerated in organs other than the lung.

This study had some limitations. First, the carcinogenic potential of the *HER2* mutation at the transmembrane domain should be confirmed in other models such as transgenic mice. Second, the rarity of these mutations in sporadic lung cancers may be the limitation for generalizability to other cases even if targeting therapies for similar types of *HER2* mutation were developed.

In conclusion, we identified a novel germline mutation in the transmembrane domain of the *HER2* in familial lung adenocarcinomas. Somatic mutation in the *HER2* transmembrane domain may be a possible cause of sporadic lung adenocarcinomas.

References

1. Bell DW, Gore I, Okimoto RA, et al. Inherited susceptibility to lung cancer may be associated with the T790M drug resistance mutation in EGFR. *Nat Genet*. 2005;37(12):1315–1316.
2. Ikeda K, Nomori H, Mori T, Sasaki J, Kobayashi T. Novel germline mutation: EGFR V843I in patient with multiple lung adenocarcinomas and family members with lung cancer. *Ann Thorac Surg*. 2008;85(4):1430–1432.
3. Ohtsuka K, Ohnishi H, Kurai D, et al. Familial lung adenocarcinoma caused by the EGFR V843I germ-line mutation. *J Clin Oncol*. 2011;29(8):e191–e192.
4. van Noesel J, van der Ven WH, van Os TA, et al. Activating germline R776H mutation in the epidermal growth factor receptor associated with lung cancer with squamous differentiation. *J Clin Oncol*. 2013;31(10):e161–e164.
5. Pao W, Girard N. New driver mutations in non-small-cell lung cancer. *Lancet Oncol*. 2011;12(2):175–180.
6. Shigematsu H, Takahashi T, Nomura M, et al. Somatic mutations of the HER2 kinase domain in lung adenocarcinomas. *Cancer Res*. 2005;65(5):1642–1646.
7. Stephens P, Hunter C, Bignell G, et al. Lung cancer: intragenic ERBB2 kinase mutations in tumours. *Nature*. 2004;431(7008):525–526.
8. Bargmann CI, Hung MC, Weinberg RA. Multiple independent activations of the neu oncogene by a point mutation altering the transmembrane domain of p185. *Cell*. 1986;45(5):649–657.
9. Bocharov EV, Mineev KS, Volynsky PE, et al. Spatial structure of the dimeric transmembrane domain of the growth factor receptor ErbB2 presumably corresponding to the receptor active state. *J Biol Chem*. 2008;283(11):6950–6956.
10. Fleishman SJ, Schlessinger J, Ben-Tal N. A putative molecular-activation switch in the transmembrane domain of erbB2. *Proc Natl Acad Sci U S A*. 2002;99(25):15937–15940.
11. Mineev KS, Bocharov EV, Pustovalova YE, Bocharova OV, Chupin VV, Arseniev AS. Spatial structure of the transmembrane domain heterodimer of ErbB1 and ErbB2 receptor tyrosine kinases. *J Mol Biol*. 2010;400(2):231–243.
12. Baselga J, Swain SM. Novel anticancer targets: revisiting ERBB2 and discovering ERBB3. *Nat Rev Cancer*. 2009;9(7):463–475.
13. Engelman JA. Targeting PI3K signalling in cancer: opportunities, challenges and limitations. *Nat Rev Cancer*. 2009;9(8):550–562.
14. Mountzios G, Planchard D, Besse B, et al. Mitogen-activated protein kinase activation in lung adenocarcinoma: a comparative study between ever smokers and never smokers. *Clin Cancer Res*. 2008;14(13):4096–4102.
15. Planchard D, Camara-Clayette V, Dorvault N, Soria JC, Fouret P. p38 Mitogen-activated protein kinase signaling, ERCC1 expression, and viability of lung cancer cells from never or light smoker patients. *Cancer*. 2012;118(20):5015–5025.

Funding

This study was supported by a Grant-in Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (25293302 to ST).

Note

H. Yamamoto, J. Soh, S. Miyoshi, and S. Toyooka conceived the project. K. Higasa, M. Sakaguchi, K. Shien, and K. Ichimura performed the experiments. H. Yamamoto, J. Soh, M. Furukawa, S. Hashida, N. Takigawa, K. Kiura, K. Tsukuda, and S. Toyooka collected the samples and assisted with the experiments. H. Yamamoto, K. Higasa, K. Shien, and K. Matsuo analyzed the data. H. Yamamoto, K. Higasa, M. Sakaguchi, F. Matsuda, and S. Toyooka prepared the manuscript with input from the other authors. S. Miyoshi, F. Matsuda, and S. Toyooka supervised the project. The authors declared no conflicts of interest.

Affiliations of authors: Department of Thoracic, Breast and Endocrinological Surgery (HY, KS, JS, MF, SH, KT, SM, ST), Department of Clinical Genomic Medicine (KS, ST), Department of Cell Biology (MS), Department of Pathology (KI), and Department of Hematology, Oncology and Respiratory Medicine (KK), Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama, Japan; Center for Genomic Medicine, Kyoto University School of Medicine, Kyoto, Japan (KH, FM); Department of General Internal Medicine 4, Kawasaki Medical School, Okayama, Japan (NT); Department of Preventive Medicine, Kyushu University Faculty of Medical Sciences, Fukuoka, Japan (KM).

Annals of human genetics

doi: 10.1111/ahg.12034

Quantitative Variation in Plasma Angiotensin-I Converting Enzyme Activity Shows Allelic Heterogeneity in the *ABO* Blood Group Locus

Chikashi Terao[1], Nervana Bayoumi[2], Colin A. McKenzie[3], Diana Zelenika[4,5], Shigeo Muro[6], Michiaki Mishima[6], The Nagahama Cohort Research Group[7], John M. C. Connell[8], Mark A. Vickers[9], G. Mark Lathrop[4,5,10], Martin Farrall[11,12], Fumihiko Matsuda[1] and Bernard D. Keavney[13,14]*

[1]*The Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan*
[2]*Physiology Department, College of Medicine, King Saud University, Riyadh, Saudi Arabia*
[3]*Tropical Metabolism Research Unit, University of the West Indies, Mona, Jamaica*
[4]*Commisariat à l'énergie Atomique (CEA), Institut Genomique, Centre National de Genotypage, Evry, France*
[5]*Fondation Jean Dausset, Centre d'Etude du Polymorphisme Humain, Paris, France*
[6]*Department of Respiratory Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan*
[7]*The following investigators belonging to Kyoto University Graduate School of Medicine (Kyoto, Japan) were core members of the Nagahama Study Group: Takeo Nakayama (Department of Health Informatics); Shinji Kosugi (Department of Medical Ethics); Akihiro Sekine, Takahisa Kawaguchi, Ryo Yamada and Yasuharu Tabara (The Center for Genomic Medicine)*
[8]*College of Medicine, Dentistry and Nursing, University of Dundee, UK*
[9]*School of Medicine, University of Aberdeen, UK*
[10]*McGill University and Genome Quebec Innovation Center, Montreal, Canada*
[11]*Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, UK*
[12]*Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK*
[13]*Institute of Genetic Medicine, Newcastle University, UK*
[14]*Institute of Cardiovascular Sciences, Manchester University, UK*

Summary

Angiotensin-I converting enzyme (ACE) occupies a pivotal role in cardiovascular homeostasis. Major loci for plasma ACE have been identified at *ACE* on Chromosome 17 and at *ABO* on Chromosome 9. We sought to characterise the genetic architecture of plasma ACE at finer resolution in two populations. We carried out a GWAS in 1810 individuals of Japanese ethnicity; this identified signals at *ACE* and *ABO* that together accounted for nearly half of the population variability of the trait. We conducted measured haplotype analysis at the *ABO* locus in 1425 members of 248 British families using haplotypes of three SNPs, which together tagged the alleles responsible for the principal blood group antigens A1, A2, B and O. Type O alleles were associated with intermediate plasma ACE activity compared to Type A1 alleles (in whom plasma ACE activity was ~36% lower) and Type B alleles (in whom plasma ACE activity was ~36% higher). We demonstrated heterogeneity among A alleles: A2 alleles were associated with plasma ACE activity that was very similar to the O alleles. Variation at *ACE* accounted for 35% of the trait variance, and variation at *ABO* accounted for 15%. A further 10% could be ascribed to polygenic effects.

Keywords: ABO blood group, angiotensin-I converting enzyme, genome wide association study, QTL

Introduction

The renin–angiotensin system plays a critical role in cardiovascular homeostasis, regulating blood pressure, arterial tone and renal salt excretion. The angiotensin I-converting enzyme (ACE) converts circulating angiotensin-I, which is

*Corresponding author: Bernard D. Keavney, Institute of Genetic Medicine, Newcastle University, Central Parkway, NE1 3BZ, UK. Tel: +44 191 241 8615; Fax: +44 191 241 8666; E-mail: bernard.keavney@newcastle.ac.uk

biologically inactive, to the active angiotensin-II and degrades the vasodilator bradykinin. ACE occupies a pivotal position in the renin–angiotensin system: drugs which inhibit ACE, or which block the cellular receptor for the angiotensin-II generated by the action of ACE, are among the most widely prescribed agents in patients with coronary artery disease, hypertension and chronic renal disease (Yusuf et al., 2000). Plasma ACE activity is related to tissue ACE activity, and is strongly influenced by genetic factors, the largest such influence being due to polymorphic variation in the *ACE* (also known as *DCP1*) gene, which encodes ACE (Rigat et al., 1990; Keavney et al., 1998). Previous segregation and linkage analysis suggested the existence of a second major quantitative trait locus (QTL) influencing plasma ACE activity (McKenzie et al., 1995). A genome wide association study (GWAS) in a population of hypertensive patients of Han Chinese ancestry subsequently found this to be located at the *ABO* gene which encodes glycosyltransferases A and B (Chung et al., 2010). Global variation in the distribution of the alleles responsible for ABO blood groups is well described; we therefore sought to confirm the identity of the second principal locus influencing ACE activity and to estimate the proportions of phenotypic variance attributable to major gene effects in two additional populations from Japan and the United Kingdom.

Methods

Study Populations and ACE Phenotyping

The discovery cohort included 1830 volunteers recruited as a part of the Nagahama Prospective Genome Cohort for Comprehensive Human Bioscience (the Nagahama Study), a community-based prospective multiomics cohort study. The study has been described in detail elsewhere (Yoshimura et al., 2012); demographics of the cohort are summarised in Table S1. In brief, a total of 9809 volunteers from Nagahama City, Shiga Prefecture, Japan, were recruited for this study from 2008 to 2010. All participants completed a detailed health questionnaire. DNA, serum and plasma samples from all participants were obtained and stored for further analysis. Samples were kept on ice immediately after they were obtained from the participants and were promptly processed. Plasma was stored at –80°C. ACE activity was quantified by the method reported by Kasahara & Ashihara (1981). Patients receiving ACE inhibitor therapy were excluded from the analyses.

The replication cohort comprised 248 British families of Northern and Western European ancestry who participated in a quantitative genetic study of cardiovascular risk factors (Palomino-Doza et al., 2008). The population collection strategy has been previously described in detail (Gaukrodger et al., 2005). In brief, families were ascertained via a hypertensive proband between 1993 and 1996, and any sibship in the family (in the generation of the proband or his/her offspring) greater than three members quantitatively assessable for blood pressure was collected. Families were extended where additional hypertensives were encountered during collection; a total of 1425 individuals participated. Families underwent detailed cardiovascular phenotyping including a questionnaire, electrocardiographic and echocardiographic measurement and measurement of 24-hour ambulatory blood pressure using an automated monitor (Keavney et al., 2000; Mayosi et al., 2008). Demographics of the cohort are presented in Table S2. Blood was drawn into multiple anticoagulants, immediately put on ice and transported rapidly to a central facility for processing. ACE activity was assessed by HPLC using a synthetic substrate, as previously described (Chiknas, 1979).

Genotyping

The 1830 volunteers in the Japanese cohort were genotyped using the Infinium Human 610-Quad Bead Chip carrying 592,044 SNP markers on a Bead Station 500G Genotyping System (Illumina, Inc., San Diego, CA, USA). There were no subjects showing call rates lower than 0.99. Kinship analysis was performed using PLINK. Of the 20 pairs of samples showing high degrees of kinship (PI-HAT > 0.4), the sample with the lower call rate in each pair was removed. 165,591 SNPs were removed either due to call rate lower than 0.95, minor allele frequency of less than 0.05, or distorted Hardy-Weinberg equilibrium ($P < 10^{-7}$). Finally, the results of 426,453 SNP markers in 1810 subjects were used for the analysis.

Three SNPs at the ABO locus (rs505922, rs8176746 and rs8176750) were typed in the entire British family cohort using matrix-assisted laser desorption/ionisation—time of flight mass spectrometry (MALDI-TOF) on a Sequenom instrument (Sequenom, San Diego, CA, USA). ABO blood group in the British families was studied in a subset of 734 individuals by multiplex polymerase chain reaction (PCR). Two pairs of primers were used to amplify exons 6 and 7 of the *ABO* gene; the amplified fragments were digested with restriction endonucleases *Hpa*II and *Kpn*I and separated by gel electrophoresis. This enabled us to call genotypes at the SNPs rs8176719, rs1053878, rs8176743 and rs8176472. As previously reported, genotypes at these SNPs, considered together, identify the A1, A2, B, O1 and O2 blood group alleles (Seltsam et al., 2003).

Statistical Methods

A quantitative linear regression analysis was first performed in the Japanese cohort to find the polymorphisms associated with ACE activity. SNP genotype imputation for SNPs within and

flanking the *ACE* (20 SNPs) and *ABO* (43 SNPs) loci was performed in the Japanese samples using the MaCH (version 1.0.10) computer program with 500 Markov sampler rounds and 200 haplotype states (Li et al., 2010). A forward-selection stepwise regression analysis was performed to identify a parsimonious subset of associated SNPs from the *ACE* and *ABO* loci in the Japanese population. This analysis was based on imputed SNP dosages using linear regression models based on marginal sums of squares and the *stepwise* procedure in Stata^{TM} v10.1 (Stata Corp, College Station, TX, USA) using a $P < 0.01$ criterion for adding SNPs to the model. Variance component proportions (R^2) were calculated from a supplementary analysis of variance based on sequential sums of squares.

Haplotyping of *ABO* in the British samples was performed using PHASE (version 2.1.1) specifying a parent-independent multiallelic model for both SNP and blood group variation. (Stephens et al., 2001; Stephens & Scheet, 2005) Pedigree analysis was performed using the Pedigree Analysis Package (PAP version 5.0) to fit maximum likelihood models including polygenic variance components (Hasstedt, 1993) to extended families; missing data is efficiently incorporated into this analysis (Elston & Stewart, 1971). For the measured haplotype analysis, the PAP quantitative major gene subroutine *qmlprmv* was modified to parametrise an additive (codominant) genetic model. Likelihoods were maximised with simultaneous estimation of haplotype frequencies assuming Hardy–Weinberg equilibrium, haplotype-specific effects on ACE activity, covariate effects, polygenic effects and residual individual-specific random (i.e. environmental) effects and estimates of standard errors were calculated with the bundled quasi-Newton nonlinear optimisation function GEMINI (Lalouel, 1979). Variance component proportions were calculated by hand using a standard additive genetic variance formula.

Results

GWAS for ACE Activity

The SNPs with the strongest association with ACE activity genome-wide are presented in Table S3. Two loci, *ACE* on Chromosome 17, and *ABO* on Chromosome 9, showed genome-wide significant association ($P < 5 \times 10^{-8}$) with plasma ACE activity (Fig. S1). SNPs mapping to the *ACE* (20 SNPs) and *ABO* (43 SNPs) loci and their immediate upstream and downstream flanking regions (50 kb, respectively) were selected for fine-mapping analysis and any missing genotype data was imputed. Stepwise linear regression then identified three SNPs with independent significant effects ($10^{-213} < P < 10^{-40}$): rs4362 at *ACE*; and rs495828 and rs8176746 at *ABO*, that together accounted for nearly half of the population variation in the trait (Table 1), with the *ACE* locus accounting for 35% and the *ABO* locus accounting for 11%.

Table 1 Forward selection stepwise regression analysis of plasma ACE activity and GWAS SNPs in the Japanese cohort. R^2 shows the proportion of variance explained by each variable.

Locus	Variable	Beta	SE	*F*-statistic	*P*-value	R^2
	Age	0.0363	0.0040	84.32	1.13E-19	0.0229
ACE	rs4362	−3.0450	0.0849	1286.30	9.60E-213	0.3494
ABO	rs495828	1.4762	0.0982	225.75	3.81E-48	0.0613
ABO	rs8176746	1.5698	0.1154	184.97	3.66E-40	0.0502

Table 2 *ABO* haplotype analysis using PHASE 2.1.1 in British families. Common haplotypes assessed in the measured haplotype analysis are shown in bold.

rs505922	rs8176746	rs8176750	Blood group allele	Conditional probability	Frequency
C	**G**	**G**	**A1**	**0.9235**	**0.1719**
C	G	G	A2	0.0049	0.0009
C	G	G	B	0.0139	0.0026
C	G	G	O1	0.0271	0.0050
C	G	G	O2	0.0307	0.0057
C	G	del	A1	0.1541	0.0096
C	**G**	**del**	**A2**	**0.8333**	**0.0517**
C	G	del	O1	0.0011	0.0001
C	G	del	O2	0.0115	0.0007
C	**T**	**G**	**B**	**1.0000**	**0.0513**
C	T	del	A2	0.2244	0.0004
C	T	del	B	0.7756	0.0014
T	G	G	A1	0.0068	0.0047
T	G	G	B	0.0055	0.0038
T	**G**	**G**	**O1**	**0.9618**	**0.6709**
T	G	G	O2	0.0259	0.0180
T	G	del	A1	0.0377	0.0000
T	G	del	O1	0.9623	0.0002
T	T	G	O1	1.0000	0.0007

Association between Haplotypes Defining Blood Groups and ACE Activity

Haplotypes of rs505922, rs8176746 and rs8176750 showed strong associations with alleles defining the different blood groups in the subset of the British families (734 samples) where blood groups were available (conditional tagging probabilities range from 0.83 to 1.00; Table 2). Haplotypes of these three SNPs that occurred at a frequency of >0.05 in the population, and accurately tagged alleles responsible for A1, A2, B and O1 phenotypes, were taken forward to the

Table 3 Measured haplotype analysis of plasma ACE activity in British families. The *ABO* haplotype is defined by rs505922, rs8176746 and rs8176750; the most strongly tagged blood group allele is shown in parentheses.

ABO haplotype (blood group allele)	Frequency	Mean	*SE*
TGG (O1)	0.647	9.295	0.084
CGG (A1)	0.212	5.989	0.231
CGdel (A2)	0.079	9.473	0.473
CTG (B)	0.063	12.808	0.528

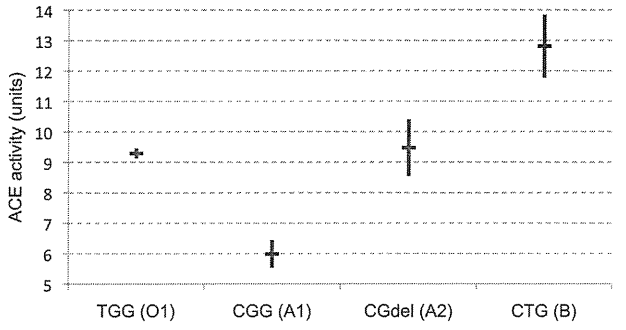

Figure 1 Measured haplotype analysis of plasma ACE activity in British families. The mean activities and 95% confidence intervals for each *ABO* haplotype are indicated by horizontal and vertical lines, respectively. *ABO* haplotypes are defined by rs505922, rs8176746 and rs8176750; the most strongly tagged blood group allele is shown in parentheses.

measured haplotype analysis of ACE activity in the total UK cohort. In the measured haplotype analyses, SNPs rs4295 and rs4392, previously shown in this cohort to tag the principal haplotype blocks influencing ACE activity at the *ACE* locus, were included as covariates. (Keavney et al., 1998) The measured haplotype analysis showed that the TGG haplotype of rs505922/rs8176746/rs8176750, which accurately tagged type O alleles and occurred at a frequency of ~65% in the population, was associated with an intermediate plasma ACE activity (Table 3). The CGG haplotype (frequency ~21%), which tagged type A1 alleles, was associated with a 36% lower plasma ACE activity than the TGG haplotype. The CTG haplotype (frequency ~6%), which tagged type B alleles, was associated with a 36% higher plasma ACE activity than the TGG haplotype. The CGdel haplotype (frequency ~8%), which tagged type A2 alleles, had a plasma ACE activity very similar to the TGG haplotype (Fig. 1). The measured haplotype analysis in the British families provided very similar estimates of variance components to the GWAS in the Japanese cohort: 35.4% of the variance was due to the *ACE* locus, and 13.0% to the *ABO* locus, with a further 9.6% attributable to polygenes and 42.1% to nongenetic residual variation.

Discussion

We have shown strong evidence for association between plasma ACE activity and genotypes at the *ABO* gene that define the major blood groups. Genome-wide analyses in a cohort of unrelated Japanese people confirmed that the strongest genetic influence on plasma ACE levels is located at the *ACE* gene itself on Chromosome 17 ($P = 1.55 \times 10^{-164}$ for rs4362) and demonstrated two further independent SNP effects at *ABO* (rs495828 and rs8176746). Stepwise regression suggested that the association at *ACE* accounted for 35% of the variability in the Japanese population and the two SNPs at *ABO* accounted for 11%. To confirm this finding, we typed SNPs at the *ABO* locus in a cohort of British families previously extensively genotypically characterised at the *ACE* locus. (McKenzie et al., 2001) Measured haplotype analysis in these families indicated that the haplotype characterizing group A1 was associated with the lowest plasma ACE level, the haplotypes characterizing groups O and A2 were associated with intermediate plasma ACE level, and the haplotype characterizing group B was associated with the highest plasma ACE level. The proportions of variance accounted for by the *ACE* and *ABO* loci in the British families were highly concordant with the Japanese cohort (35.4% and 13.0%, respectively) with an additional 9.6% attributable to polygenic effects. The *ACE* and *ABO* loci make the most substantial contribution to population variance in plasma ACE levels; also, there is appreciable heterogeneity in plasma ACE levels among the alleles specifying the two subgroups A1 and A2.

The *ABO* gene encodes a glycosyltransferase. Genetic variation in *ABO* results in the production of two differently named protein products: glycosyltransferase-A and glycosyltransferase-B. Glycosyltransferase-A transfers N-acetylgalactosamine to an acceptor glycoconjugate on the glycosphingolipid H-antigen, which is strongly present on the surface of red blood cells and more weakly present on a wide range of other cell types. Glycosyltransferase-B transfers D-galactose to the same position on the H-antigen. These glycosyltransferase activities define the blood group antigens A and B, respectively. AB heterozygotes have molecules with both A and B antigens present on the red cell surface. Mutations which inactivate the glycosyltransferase encoded by *ABO* result in nonmodification of the H-antigen, which characterises blood group O. A and B alleles are both dominant to O.

Previous studies have shown association between either *ABO* genotypes or ABO blood groups and plasma ACE activity. Cidl et al. found blood groups A and O to have similar levels of plasma ACE activity and groups AB and B to have progressively higher levels among 197 Caucasian subjects of Eastern European origin. (Cidl et al., 1996) Chung et al. performed a GWAS for plasma ACE activity among 1023 subjects with young-onset hypertension, replicating their findings in a

study of 428 hypertension pedigrees, all of self-reported Han Chinese ethnicity (Chung et al., 2010). By contrast with the findings of Cidl et al., these authors found plasma ACE activity in blood group A individuals to be 86% of that in the reference group O, and plasma ACE activity in blood group B individuals to be 114% of that in the reference group O. However, neither of these previous studies carried out measured haplotype analyses to determine the effect of the combination of SNPs defining the principal blood group antigens on plasma ACE levels, and neither conducted formal two-locus genetic analyses to determine the relative contribution of *ABO* haplotypes and *ACE* haplotypes on phenotypic variance. Moreover, the reason for the discordance between the two reports with respect to the relative levels of plasma ACE activity in type A and type O individuals remained unclear. Our findings illustrate significant differences in plasma ACE activity associated with the haplotypes defining the type A subgroups A1 and A2. The A1-defining haplotype had ~60% of the plasma ACE activity of the O-defining haplotype, whereas the plasma ACE activity of the A2-defining haplotype was not significantly different from the O-defining haplotype. In common with the previous studies, we found that the B-defining haplotype was associated with a ~40% higher plasma ACE activity than the O-defining haplotype.

The effects of the different *ABO* haplotypes on plasma ACE activity were assessed in a codominant (allelic association) model. This may explain the substantially smaller proportion of variation that could be accounted for by ABO blood grouping in the study of Cidl et al. (1996), when compared to the present study (since A and B alleles are dominant to O with respect to the determination of blood group). Differences in the proportions of A1 and A2 alleles in the previously studied populations may also contribute to the discrepancy in the results of previous studies: among Asian populations such as those studied by Chung et al. (2010), the A2 allele is extremely rare, while among European populations it comprises up to 25% of all A alleles. The A2 allele shows a weaker activity than the A1 allele in adding N-acetylgalactosamine to the acceptor glycoconjugate of the H antigen, which is present on red blood cells. At the DNA level, the A2 allele differs from the A1 in two ways: a single amino acid change (Pro156Leu) and a 1061delC mutation causing a frameshift that extends the reading frame by 64 nucleotides (Yamamoto et al., 1992).

Recently, several significant associations between the *ABO* locus and complex diseases have been reported. Association between ABO blood group and the risk of myocardial infarction (MI) was first observed decades ago; a GWAS approach recently confirmed strong association between non-O blood group and the risk of MI, although not of coronary artery disease without MI (Reilly et al., 2011). The likely principal mechanism is a differential susceptibility to thrombosis: people of blood group O have 25% lower plasma levels of von Willebrand factor and Factor VIII, both members of the blood coagulation cascade, and SNPs defining group O are associated with less coagulable blood on standard laboratory coagulation tests. (Tang et al., 2012) Congruent with this observation, *ABO* genotypes are the strongest risk factor genome-wide for venous thromboembolism (Qi et al., 2010). Nevertheless, other potential mechanisms whereby the pleiotropic *ABO* locus could contribute to MI risk have been identified through GWAS approaches. *ABO* genotypes are associated with plasma levels of lipids, and of other plasma factors associated with coronary artery disease including soluble ICAM-1, E-selectin and P-selectin (Barbalic et al., 2010; Qi et al., 2010; Pare et al., 2011). GWAS studies have also identified SNPs at *ABO* responsible for the long-standing epidemiological observations of association between blood group O and higher risk of duodenal ulceration; (Tanikawa et al. 2012) and lower risks of gastric and pancreatic cancer (Amundadottir et al., 2009). Plasma ACE activity has not, in general, been systematically studied as a risk factor for any of the conditions associated with ABO blood groups. Since the *ABO* haplotype defining type O in our study was associated with intermediate levels of plasma ACE activity compared with types A and B, it is unlikely that plasma ACE activity lies in the causal pathway for any of these conditions, in which type O lies at the upper or lower extreme of the risk distribution. Further research will be required to define the mechanism whereby *ABO* haplotypes are associated with plasma ACE activity, which remains uncertain. A plausible hypothesis is that it may involve differential affinities of receptors involved in the clearance of ACE, a heavily glycosylated protein, for the A B and H glycosylation motifs that depend upon *ABO* genotype.

Although evidence remains inconclusive, it has been suggested that blood group O protects against severe malaria, and that this may have constituted a selective advantage favouring the alleles responsible for the group O phenotype throughout human evolutionary history (Anstee, 2010). In this regard it is of interest that angiotensin-II has recently been shown to reduce erythrocyte invasion *in vitro* by *P. falciparum* in a dose-dependent manner, and that the deletion allele of the *ACE* insertion/deletion polymorphism, which is associated with higher plasma ACE activity, has been reported to be protective against cerebral malaria (Dhangadamajhi et al., 2010; Saraiva et al., 2011). Further research into the question of whether the major genetic effects on plasma ACE activity at the *ACE* and *ABO* loci contribute significantly to differential susceptibility to severe malaria would be of interest.

In conclusion, this study has confirmed the existence of two major genetic loci influencing plasma levels of ACE, and quantified the contribution of the two loci to the population variability of the trait in two populations. We have also demonstrated previously unrecognised heterogeneity among

alleles responsible for group A antigens with respect to their effect on plasma ACE.

Acknowledgements

The principal acknowledgement is to the participants in the Japanese and British cohorts. We are grateful to the Nagahama City Office and the nonprofit organisation "Zeoji Club" for their help in performing the Nagahama Study. The Nagahama Study is in part supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan; and by the Takeda Science Foundation. The collection and genotyping of the British families was funded by the Wellcome Trust and the British Heart Foundation. BK holds a British Heart Foundation Personal Chair. MF acknowledges support from the Wellcome Trust core award (090532/Z/09/Z) and the BHF Centre of Clinical Research Excellence in Oxford. ML acknowledges support from Genome Québec, le Ministère de l'Enseignement supérieur, de la Recherche, de la Science et de la Technologie (MESRST) Québec and McGill University, and from the Agence Nationale de la Recherche (ANR), France, for the Labex project "Medical Genomics." No author has a conflict of interest to declare in respect of this work.

References

Amundadottir, L., Kraft, P., Stolzenberg-Solomon, R. Z., Fuchs, C. S., Petersen, G. M., Arslan, A. A., Bueno-De-Mesquita, H. B., Gross, M., Helzlsouer, K., Jacobs, E. J., Lacroix, A., Zheng, W., Albanes, D., Bamlet, W., Berg, C. D., Berrino, F., Bingham, S., Buring, J. E., Bracci, P. M., Canzian, F., Clavel-Chapelon, F., Clipp, S., Cotterchio, M., De Andrade, M., Duell, E. J., Fox, J. W., Jr., Gallinger, S., Gaziano, J. M., Giovannucci, E. L., Goggins, M., Gonzalez, C. A., Hallmans, G., Hankinson, S. E., Hassan, M., Holly, E. A., Hunter, D. J., Hutchinson, A., Jackson, R., Jacobs, K. B., Jenab, M., Kaaks, R., Klein, A. P., Kooperberg, C., Kurtz, R. C., Li, D., Lynch, S. M., Mandelson, M., Mcwilliams, R. R., Mendelsohn, J. B., Michaud, D. S., Olson, S. H., Overvad, K., Patel, A. V., Peeters, P. H., Rajkovic, A., Riboli, E., Risch, H. A., Shu, X. O., Thomas, G., Tobias, G. S., Trichopoulos, D., Van Den Eeden, S. K., Virtamo, J., Wactawski-Wende, J., Wolpin, B. M., Yu, H., Yu, K., Zeleniuch-Jacquotte, A., Chanock, S. J., Hartge, P., & Hoover, R. N. (2009) Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. *Nat Genet* **41**, 986–990.

Anstee, D. J. (2010) The relationship between blood groups and disease. *Blood* **115**, 4635–4643.

Barbalic, M., Dupuis, J., Dehghan, A., Bis, J. C., Hoogeveen, R. C., Schnabel, R. B., Nambi, V., Bretler, M., Smith, N. L., Peters, A., Lu, C., Tracy, R. P., Aleksic, N., Heeriga, J., Keaney, J. F., Jr., Rice, K., Lip, G. Y., Vasan, R. S., Glazer, N. L., Larson, M. G., Uitterlinden, A. G., Yamamoto, J., Durda, P., Haritunians, T., Psaty, B. M., Boerwinkle, E., Hofman, A., Koenig, W., Jenny, N. S., Witteman, J. C., Ballantyne, C., & Benjamin, E. J. (2010) Large-scale genomic studies reveal central role of ABO in sP-selectin and sICAM-1 levels. *Hum Mol Genet* **19**, 1863–1872.

Chiknas, S. G. (1979) A liquid chromatography-assisted assay for angiotensin-converting enzyme (peptidyl dipeptidase) in serum. *Clin Chem* **25**, 1259–1262.

Chung, C. M., Wang, R. Y., Chen, J. W., Fann, C. S., Leu, H. B., Ho, H. Y., Ting, C. T., Lin, T. H., Sheu, S. H., Tsai, W. C., Chen, J. H., Jong, Y. S., Lin, S. J., Chen, Y. T., & Pan, W. H. (2010) A genome-wide association study identifies new loci for ACE activity: potential implications for response to ACE inhibitor. *Pharmacogenomics J* **10**, 537–544.

Cidl, K., Strelcova, L., Znojil, V., & Vachi, J. (1996) Angiotensin I-converting enzyme (ACE) polymorphism and ABO blood groups as factors codetermining plasma ACE activity. *Exp Hematol* **24**, 790–794.

Dhangadamajhi, G., Mohapatra, B. N., Kar, S. K., & Ranjit, M. (2010) Gene polymorphisms in angiotensin I converting enzyme (ACE I/D) and angiotensin II converting enzyme (ACE2 C–>T) protect against cerebral malaria in Indian adults. *Infect Genet Evol* **10**, 337–341.

Elston, R. C., & Stewart, J. (1971) A general model for the genetic analysis of pedigree data. *Hum Hered* **21**, 523–542.

Gaukrodger, N., Mayosi, B. M., Imrie, H., Avery, P., Baker, M., Connell, J. M., Watkins, H., Farrall, M., & Keavney, B. (2005) A rare variant of the leptin gene has large effects on blood pressure and carotid intima-medial thickness: a study of 1428 individuals in 248 families. *J Med Genet* **42**, 474–478.

Hasstedt, S. J. (1993) Variance components/major locus likelihood approximation for quantitative, polychotomous, and multivariate data. *Genet Epidemiol* **10**, 145–158.

Kasahara, Y., & Ashihara, Y. (1981) Colorimetry of angiotensin-I converting enzyme activity in serum. *Clin Chem* **27**, 1922–1925.

Keavney, B., Bird, R., Caiazza, A., Casadei, B., & Conway, J. (2000) Measurement of blood pressure using the auscultatory and oscillometric methods in the same cuff deflation: validation and field trial of the A&D TM2421 monitor. *J Hum Hypertens* **14**, 573–579.

Keavney, B., Mckenzie, C. A., Connell, J. M., Julier, C., Ratcliffe, P. J., Sobel, E., Lathrop, M., & Farrall, M. (1998) Measured haplotype analysis of the angiotensin-I converting enzyme gene. *Hum Mol Genet* **7**, 1745–1751.

Lalouel, J.-M. (1979) GEMINI: A computer program for optimization of general nonlinear functions. Technical Reports. Honolulu.

Li, Y., Willer, C. J., Ding, J., Scheet, P., & Abecasis, G. R. (2010) MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* **34**, 816–834.

Mayosi, B. M., Avery, P. J., Farrall, M., Keavney, B., & Watkins, H. (2008) Genome-wide linkage analysis of electrocardiographic and echocardiographic left ventricular hypertrophy in families with hypertension. *Eur Heart J* **29**, 525–530.

McKenzie, C. A., Abecasis, G. R., Keavney, B., Forrester, T., Ratcliffe, P. J., Julier, C., Connell, J. M., Bennett, F., Mcfarlane-Anderson, N., Lathrop, G. M., & Cardon, L. R. (2001) Trans-ethnic fine mapping of a quantitative trait locus for circulating angiotensin I-converting enzyme (ACE). *Hum Mol Genet* **10**, 1077–1084.

McKenzie, C. A., Julier, C., Forrester, T., Mcfarlane-Anderson, N., Keavney, B., Lathrop, G. M., Ratcliffe, P. J., & Farrall, M. (1995) Segregation and linkage analysis of serum angiotensin I-converting enzyme levels: evidence for two quantitative-trait loci. *Am J Hum Genet* **57**, 1426–1435.

Palomino-Doza, J., Rahman, T. J., Avery, P. J., Mayosi, B. M., Farrall, M., Watkins, H., Edwards, C. R., & Keavney, B. (2008) Ambulatory blood pressure is associated with polymorphic variation in P2X receptor genes. *Hypertension* **52**, 980–985.

Pare, G., Ridker, P. M., Rose, L., Barbalic, M., Dupuis, J., Dehghan, A., Bis, J. C., Benjamin, E. J., Shiffman, D., Parker, A. N., & Chasman, D. I. (2011) Genome-wide association analysis of soluble ICAM-1 concentration reveals novel associations at the NFKBIK, PNPLA3, RELA, and SH2B3 loci. *PLoS Genet* **7**, e1001374.

Qi, L., Cornelis, M. C., Kraft, P., Jensen, M., Van Dam, R. M., Sun, Q., Girman, C. J., Laurie, C. C., Mirel, D. B., Hunter, D. J., Rimm, E., & Hu, F. B. (2010) Genetic variants in ABO blood group region, plasma soluble E-selectin levels and risk of type 2 diabetes. *Hum Mol Genet* **19**, 1856–1862.

Reilly, M. P., Li, M., He, J., Ferguson, J. F., Stylianou, I. M., Mehta, N. N., Burnett, M. S., Devaney, J. M., Knouff, C. W., Thompson, J. R., Horne, B. D., Stewart, A. F., Assimes, T. L., Wild, P. S., Allayee, H., Nitschke, P. L., Patel, R. S., Martinelli, N., Girelli, D., Quyyumi, A. A., Anderson, J. L., Erdmann, J., Hall, A. S., Schunkert, H., Quertermous, T., Blankenberg, S., Hazen, S. L., Roberts, R., Kathiresan, S., Samani, N. J., Epstein, S. E., & Rader, D. J. (2011) Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. *Lancet* **377**, 383–392.

Rigat, B., Hubert, C., Alhenc-Gelas, F., Cambien, F., Corvol, P., & Soubrier, F. (1990) An insertion/deletion polymorphism in the angiotensin I-converting enzyme gene accounting for half the variance of serum enzyme levels. *J Clin Invest* **86**, 1343–1346.

Saraiva, V. B., De Souza Silva, L., Ferreira-Dasilva, C. T., Da Silva-Filho, J. L., Teixeira-Ferreira, A., Perales, J., Souza, M. C., Henriques, M., Caruso-Neves, C., & De Sa Pinheiro, A. A. (2011) Impairment of the Plasmodium falciparum erythrocytic cycle induced by angiotensin peptides. *PLoS One* **6**, e17174.

Seltsam, A., Hallensleben, M., Kollmann, A., & Blasczyk, R. (2003) The nature of diversity and diversification at the ABO locus. *Blood* **102**, 3035–42.

Stephens, M., & Scheet, P. (2005) Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am J Hum Genet* **76**, 449–462.

Stephens, M., Smith, N. J., & Donnelly, P. (2001) A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* **68**, 978–989.

Tang, W., Schwienbacher, C., Lopez, L. M., Ben-Shlomo, Y., Oudot-Mellakh, T., Johnson, A. D., Samani, N. J., Basu, S., Gogele, M., Davies, G., Lowe, G. D., Tregouet, D. A., Tan, A., Pankow, J. S., Tenesa, A., Levy, D., Volpato, C. B., Rumley, A., Gow, A. J., Minelli, C., Yarnell, J. W., Porteous, D. J., Starr, J. M., Gallacher, J., Boerwinkle, E., Visscher, P. M., Pramstaller, P. P., Cushman, M., Emilsson, V., Plump, A. S., Matijevic, N., Morange, P. E., Deary, I. J., Hicks, A. A., & Folsom, A. R. (2012) Genetic associations for activated partial thromboplastin time and prothrombin time, their gene expression profiles, and risk of coronary artery disease. *Am J Hum Genet* **91**, 152–162.

Tanikawa, C., Urabe, Y., Matsuo, K., Kubo, M., Takahashi, A., Ito, H., Tajima, K., Kamatani, N., Nakamura, Y., & Matsuda, K. (2012) A genome-wide association study identifies two susceptibility loci for duodenal ulcer in the Japanese population. *Nat Genet* **44**, 430–434.

Yamamoto, F., Mcneill, P. D., & Hakomori, S. (1992) Human histo-blood group A2 transferase coded by A2 allele, one of the A subtypes, is characterized by a single base deletion in the coding sequence, which results in an additional domain at the carboxyl terminal. *Biochem Biophys Res Commun* **187**, 366–374.

Yoshimura, K., Nakayama, T., Sekine, A., Matsuda, F., Kosugi, S., Yamada, R., Shimizu, Y., Kanematsu, A., & Ogawa, O. (2012) B-type natriuretic peptide as an independent correlate of nocturnal voiding in Japanese women. *Neurourol Urodyn* **31**, 1266–1271.

Yusuf, S., Sleight, P., Pogue, J., Bosch, J., Davies, R., Dagenais, G., & The Heart Outcomes Prevention Evaluation Study Investigators. (2000) Effects of an angiotensin-converting-enzyme inhibitor, ramipril, on cardiovascular events in high-risk patients. *N Engl J Med* **342**, 145–153.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1 Demographics of the Japanese discovery cohort.

Table S2 Demographics of the British replication cohort.

Table S3 Single-SNP analyses of GWAS SNPs at *ACE* and *ABO* loci in the Japanese cohort.

Figure S1 Manhattan plot of GWAS for plasma ACE in Japanese cohort, showing significant evidence of association at the *ACE* locus on Chromosome 17 and the *ABO* locus on Chromosome 9.

Received: 26 February 2013
Accepted: 16 May 2013

PLOS ONE

Three Groups in the 28 Joints for Rheumatoid Arthritis Synovitis – Analysis Using More than 17,000 Assessments in the KURAMA Database

Chikashi Terao[1,2]*, Motomu Hashimoto[2,3], Keiichi Yamamoto[4], Kosaku Murakami[2], Koichiro Ohmura[2], Ran Nakashima[2], Noriyuki Yamakawa[2], Hajime Yoshifuji[2], Naoichiro Yukawa[2], Daisuke Kawabata[2], Takashi Usui[2], Hiroyuki Yoshitomi[5], Moritoshi Furu[3,5], Ryo Yamada[1,6], Fumihiko Matsuda[1,7,8], Hiromu Ito[3,5], Takao Fujii[2,3], Tsuneyo Mimori[2,3]

1 Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan, **2** Department of Rheumatology and Clinical Immunology, Kyoto University Graduate School of Medicine, Kyoto, Japan, **3** Department of the Control for Rheumatic Diseases, Kyoto University Graduate School of Medicine, Kyoto, Japan, **4** Department of Clinical Trial Design and Management, Translational Research Center, Kyoto University Hospital, Kyoto, Japan, **5** Department of Orthopaedic Surgery, Kyoto University Graduate School of Medicine, Kyoto, Japan, **6** Unit of Statistical Genetics Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan, **7** Institut National de la Sante et de la Recherche Medicale (INSERM) Unite U852, Kyoto University Graduate School of Medicine, Kyoto, Japan, **8** CREST Program, Japan Science and Technology Agency, Kawaguchi, Saitama, Japan

Abstract

Rheumatoid arthritis (RA) is a joint-destructive autoimmune disease. Three composite indices evaluating the same 28 joints are commonly used for the evaluation of RA activity. However, the relationship between, and the frequency of, the joint involvements are still not fully understood. Here, we obtained and analyzed 17,311 assessments for 28 joints in 1,314 patients with RA from 2005 to 2011 from electronic clinical chart templates stored in the KURAMA (Kyoto University Rheumatoid Arthritis Management Alliance) database. Affected rates for swelling and tenderness were assessed for each of the 28 joints and compared between two different sets of RA patients. Correlations of joint symptoms were analyzed for swellings and tenderness using kappa coefficient and eigen vectors by principal component analysis. As a result, we found that joint affected rates greatly varied from joint to joint both for tenderness and swelling for the two sets. Right wrist joint is the most affected joint of the 28 joints. Tenderness and swellings are well correlated in the same joints except for the shoulder joints. Patients with RA tended to demonstrate right-dominant joint involvement and joint destruction. We also found that RA synovitis could be classified into three categories of joints in the correlation analyses: large joints with wrist joints, PIP joints, and MCP joints. Clustering analysis based on distribution of synovitis revealed that patients with RA could be classified into six subgroups. We confirmed the symmetric joint involvement in RA. Our results suggested that RA synovitis can be classified into subgroups and that several different mechanisms may underlie the pathophysiology in RA synovitis.

Citation: Terao C, Hashimoto M, Yamamoto K, Murakami K, Ohmura K, et al. (2013) Three Groups in the 28 Joints for Rheumatoid Arthritis Synovitis – Analysis Using More than 17,000 Assessments in the KURAMA Database. PLoS ONE 8(3): e59341. doi:10.1371/journal.pone.0059341

Editor: Bernhard Kaltenboeck, Auburn University, United States of America

Received October 21, 2012; **Accepted** February 12, 2013; **Published** March 12, 2013

Funding: This study was supported by research grants from Mitsubishi Tanabe Pharma Corporation (http://www.mt-pharma.co.jp/e/), Eisai Co., Ltd. (http://www.eisai.co.jp/index.html), Abbott Japan Co., Ltd. (http://www.abbott.co.jp/), Chugai Pharmaceutical Co., Ltd. (http://www.chugai-pharm.co.jp/hc/ss/english/index.html), Pfizer Japan Inc. (http://www.pfizer.co.jp/pfizer/english/company/), and Bristol-Myyers K.K. (http://www.bms.co.jp/). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. No additional external funding was received for this study.

Competing Interests: The KURAMA database was supported by funding from Mitsubishi Tanabe Pharma Corporation, Eisai Co., Ltd., Abbott Japan Co., Ltd., Chugai Pharmaceutical Co., Ltd., Pfizer Japan Inc. and Bristol-Myyers. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: a0001101@kuhp.kyoto-u.ac.jp

Introduction

Rheumatoid arthritis (RA) is the most frequent inflammatory arthritis worldwide affecting 0.5 to 1% of the population [1]. As RA is a bone-destructive disease and functional impairment caused by joint damage is well correlated with swelling and tenderness of joints [2–3], the evaluation of joints in patients with RA is very important to assess disease activity and predict the risk of future joint deformity. ACR core set [4] and DAS (disease activity score) [5–6] were developed for evaluation of disease activity in RA. Recently, the three composite indices, namely, DAS28 [5], simplified disease activity index (SDAI) [7] and clinical disease activity index (CDAI) [8] are frequently used for disease activity evaluation among rheumatologists. All of the three indices are shown to be well correlated with future joint destruction [7,9]. These three methods include the same 28 joints for evaluation of disease activity, namely, bilateral wrist, 1st to 5th metacarpal (MCP) joints and proximal interphalangeal (PIP) joints, elbow, shoulder, and knee joints. Though RA is known to show symmetric joint symptoms [10], the frequency of bilateral joint symptoms and the correlations between each joint symptom are not fully analyzed by using large numbers of joint assessments. There are several reports of successful prediction of joint damage using a reduced number of joints for evaluation by ultrasonogra-

phy [11–12]. These reports raise the possibility that some of the 28 joints are less frequently involved, and are less informative for disease activity. Analyses for characterization of joint symptoms would uncover correlations of unexpected joint symptoms and distribution of synovitis in RA.

Here, we analyzed the distribution of affected joints in the 28 joints in patients with RA using more than 17,000 joint assessments from 1,314 patients with RA and showed that synovitis in RA patients can be classified into three groups. We also showed that affected rates of the 28 joints greatly vary in RA patients, and that RA patients could be classified into subgroups based on the distribution of joint synovitis.

Results

Frequency order of joints involvement

We recruited 17,311 assessments for the 28 joints in 1,314 patients with RA from 2005 to 2011. A summary of the registered patients is listed in Table 1. The distribution of the number of patients with RA in each year and the number of joint assessments for each patient are shown in Figure S1. We analyzed how often each of the 28 joints was tender or swollen in patients with RA in 2011. From the analysis of 735 patients, we found that the frequency of joint swelling and tenderness in the 28 joints is widely different from joint to joint (Figure 1 and Table S1). The wrist joints were the most frequently affected joints for swelling and tenderness. The frequency of the right wrist joint being affected was more than four times as high as the least frequently affected joint. Many of the joints showed right-dominant tenderness (eleven of fourteen joints, $p = 0.057$, binomial test), indicating mostly right-handedness. We found strong correlations for the affected rates of each joint between swellings and tenderness except for shoulder joints (Spearman's rank-sum coefficient, $rho = 0.70$ and $p = 3.8 \times 10^{-5}$, Figure 1, Table S1). Shoulder joints showed much higher frequencies of tenderness than those of swellings.

Next, we tried to replicate the order of affected frequencies of the 28 joints and the correlation between tenderness and swellings in different RA patients. We obtained 579 patients whose joints data were not available for 2011, indicating we analyzed independent RA patients. We found that the order of the affected joint frequencies were well correlated for both swelling and tenderness among different sets of RA patients (Spearman's rank-sum coefficient, rho:0.815 and 0.904, $p = 1.3 \times 10^{-7}$ and $p = 4.6 \times 10^{-11}$ for swelling and tenderness, respectively, Figure S2). We also confirmed that rates of tenderness were well correlated with those of swellings in the 28 joints in the 579 patients (rho:0.604). These results indicate that some of the 28 joints are more likely to develop arthritis than the others in RA patients. The swelling and tenderness correlate with each other except for shoulder joints.

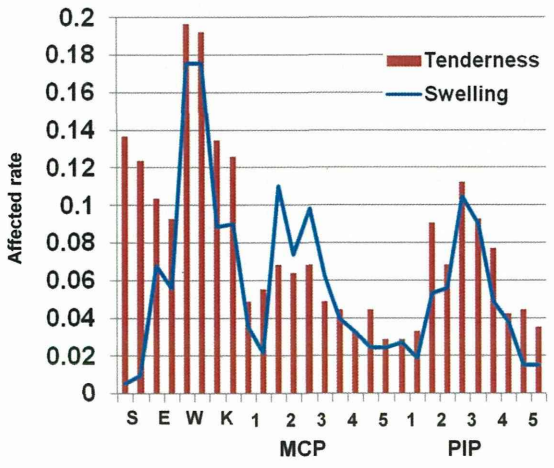

Figure 1. Affected rate of joint symptoms. Affected rate of joint symptoms. Each joint is arranged in the order of right and left. S:shoulder, E:elbow, W:wrist, K:knee.
doi:10.1371/journal.pone.0059341.g001

Table 1. Summary of the KURAMA database.

	The KURAMA database
RA patients	1314
Age (mean±SD)	60.2±15.1
female ratio	81.70%
disease duration (years)	12.2±9.8
Stage*	2.75±1.17
Class*	1.87±0.69

*Stage and Class represent Steinbrocker's stage and class, respectively.
SD: standard deviation.
doi:10.1371/journal.pone.0059341.t001

Whether the right-dominant involvement of joints in patients with RA is associated with joint destruction was analyzed. Joint destruction in the hand was evaluated for 246 patients with RA by modified Sharp score [13]. The six elements of the scores were separately analyzed, namely erosion of PIP, MCP, and wrist joints (we defined as joints other than MCP and PIP in hand) and narrowing of PIP, MCP, and wrist joints. We found that five out of six elements showed right-dominant destruction. In particular, narrowing and erosion of MCP joints showed a statistically significant right-dominance in binomial test ($p <= 0.0050$, Table S2).

Three groups of 28 joints in RA synovitis

Next we analyzed correlations of joint symptoms between the 28 joints. We randomly picked up one assessment from each of the 1,314 patients to maximize the power. When the correlation of tenderness of the 28 joints was analyzed with kappa coefficient, we confirmed that each joint showed a symmetric involvement (Figure 2A). The results also showed that the tenderness of large joints and wrist joints are not correlated with the tenderness of PIP and MCP joints. We found that the tenderness of MCP joints was especially well correlated with each other and that PIP joints tenderness was well correlated with each other. The correlation of swelling in the 28 joints showed the same tendency as that of tenderness, namely, symmetric joint involvement, correlations between large joints and wrist joints, and no strong correlations between wrist joints and other small joints (Figure 2B).

Next we used eigen vectors of principal component analysis to assess the correlations of the 28 joints involvement. When we analyzed correlations of tenderness, eigen vectors revealed that PIP and MCP joints can be clearly distinguished from large joints and wrist joints (Figure 3A). PIP joints and MCP joints turned out to make independent groups after excluding large joints and wrist joints (Figure 3B). These three groups of affected joints were found both for tenderness and swelling (Figure 3C and 3D). We confirmed these three correlation groups in four independent resampling analyses by randomly picking up one assessment from each of the 1,314 patients four times (data not shown). The three groups were observed in the two independent sets of RA patients which were used in the analysis of joints involvement frequency

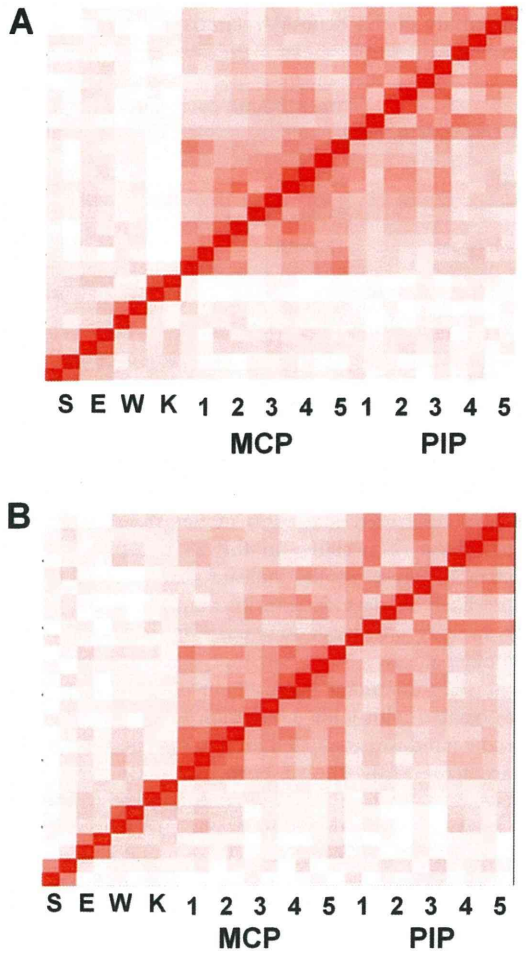

Figure 2. Correlations between the 28 joint symptoms. Brightness of the red color corresponds to the strength of correlations between joint tenderness (A) or swellings (B), using the Kappa coefficient. Each joint is arranged in the order of right and left. The joint order in the y axis is the same as the x axis. The result is a representative of five analyses based on resampled assessments. S:shoulder, E:elbow, W:wrist, K:knee.
doi:10.1371/journal.pone.0059341.g002

(Figure S3). In addition, no significant difference was observed in the relationship of the three groups of joint involvement when we divided the 1,314 patients into two groups according to the patients' caring physicians (Figure S4). We confirmed the three groups by resampling four times for each analysis (data not shown). These results indicate that these three groups were not due to specific patients, examiners, or time of evaluation.

Taken together, the correlation analyses using kappa coefficient and eigen vectors in principal component analysis indicated that there are three correlated groups of joints in RA synovitis, namely, large joints with wrist joints (which we express as "large and wrist joints"), PIP joints, and MCP joints.

Subgroups of patients with RA

We performed a clustering analysis of 5,383 evaluations of 28 joints from 1,314 patients with RA. Six subgroups of evaluations of 28 joints were observed (Figure 4). Each of the subgroups was characterized by 1) no synovitis (34.6%), 2) mild activity with dominant involvement of large and wrist joints (17.4%), 3) dominant involvement of MCP joints (18.3%), 4) dominant involvement of PIP joints (9.3%), 5) active synovitis (4.1%), and 6) moderate activity with dominant involvement of large and wrist joints (16.4%) (Table S3). Whether patients with RA are classified into the same subgroups was analyzed. There were 998 patients with four or five evaluations, and of these, 734 were categorized into the regular groups across different evaluations, indicating that the patterns of synovitis in the same patients were stable. Analysis of joint destruction in each subgroup revealed that the sixth subgroup demonstrated dominant destruction of large and wrist joints compared with MCP and PIP joints ($p <= 2.8 \times 10^{-5}$, Figure S5 and Figure S6).

Discussion

Since RA is a joint destructive autoimmune arthritis and joint damage occurs rapidly in the early stages of the disease course [14], the development of a quantitative scale which assesses disease activity and predicts joint damage is very important. After DAS and ACR core sets were introduced, DAS28, SDAI, and CDAI were developed to evaluate disease activity and easily calculate the disease activity score in patients with RA. All three indices were shown to be well correlated with future joint destruction and they share the same 28 joints for evaluation. Joint symptoms especially joint swelling is known to correlate with future joint damage [3]. While these indices were developed for use in clinical trials such as responsiveness to treatment, they are used by rheumatologists in daily clinical practice and they are reported to coincide very well among different examiners [9]. Characterizing the relative affected frequency of each joint and analysis of correlation between joint symptoms are important to analyze the basic mechanisms of synovitis and to efficiently select the joints to predict future joint destruction. However, there is no detailed analysis to address the correlations between the 28-joint symptoms.

In the current study, we characterized the 28-joint symptoms using large numbers of joint assessments. While we reported the affected rates of each joint in the 28 joints for tenderness and swelling of RA patients registered in the KURAMA database in 2011 as a representative (Table S1), these rates should not be generalized considering large effects of treatment especially biologics agents on joint symptoms. Thus, we focused on relative frequencies of joint involvement for the 28 joints. The affected frequency pattern was compared between the two sets of RA patients, and there were no apparent differences between the two sets for both tenderness and swelling. We also showed that joint symptoms in RA could be classified into three groups both for tenderness and swelling. Our analysis also demonstrated that patients with RA can be regularly classified into six subgroups based on patterns of joint symptoms. These results suggest that regular RA joint involvement pattern, including relative frequency and groups of joints, is largely maintained in RA patients. In addition, we confirmed that these patterns of joint involvement were not attributed to evaluators and fractions of RA patients.

It is interesting that the affected frequencies greatly varied from joint to joint, and the rate of the most highly affected joint was more than four times as high as the least-affected joint. The affected frequencies indicated that wrist joints were the most frequently affected. It should be noted that surface area may have influenced the sensitivity of detecting synovitis in physical exams when different joints were compared. The relatively high frequency of tenderness and swelling in large and wrist joints compared with MCP and PIP joints can be explained by this difference in surface area. However, surface area cannot fully explain the highest frequency of wrist involvement and different frequencies within the MCP or PIP joints. A dominant involve-

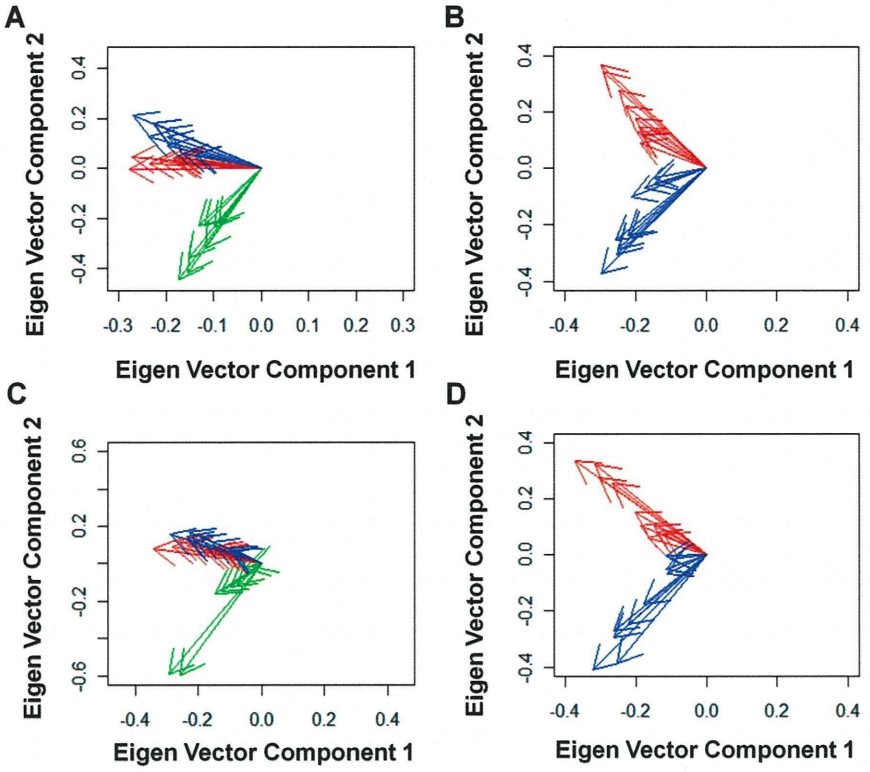

Figure 3. Relationship of the 28-joint involvement. The 1st and 2nd components of eigen vectors of the joint symptoms are plotted, using principal component analysis of the 28 joint involvement for tenderness (A) and swelling (C) or using that of the 20 joint involvement other than large and wrist joints for tenderness (B) and swelling (D). The results are representatives of five analyses based on resampled assessments. Green: large and wrist joints. Red: MCP joints. Blue: PIP joints.
doi:10.1371/journal.pone.0059341.g003

ment of right joints seemed to indicate a majority of the study population being right-handed in spite of the small difference of affected rates between bilateral joints. We also demonstrated that the right dominant involvement was also true for joint destruction. We could not compare the joint involvement and joint destruction between right-handed patients and left-handed patients due to a lack of information regarding handedness of patients.

Correlation analysis confirmed the well-known symmetric joint involvement in patients with RA. Strong correlations of tenderness and swelling in the same joints except for shoulder joints may indicate low sensitivity of shoulder swelling in the physical exams and common mechanisms of swelling and tenderness. It is striking that joint symptoms can be classified into three groups based on correlation analysis and principal component analysis. The association observed between the symptoms in the wrist joints and the large joints is worth noting, since wrist joints are regarded as small joints according to ACR/EULAR criteria set in 2010. As wrist joints are much closer to other small joints than large joints, the relationship between wrist joints and large joints cannot be explained by the distance of joints. The distance of joints cannot explain the two different groups of MCP and PIP joints either. While symptoms of large and wrist joints are not related with those of MCP and PIP joints, they were not very strongly correlated with each other, compared with correlations among PIP joints or MCP joints. This may indicate that there are no common strong factors which predispose large and wrist joints to swelling and tenderness in patients with RA.

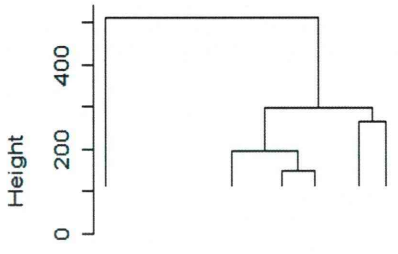

Figure 4. Six subgroups of evaluations of the 28 joints in RA. Results of clustering analysis with Ward method using randomly obtained 5,383 evaluations of the 28 joints in 1,314 patients were plotted.
doi:10.1371/journal.pone.0059341.g004

We also showed that patients with RA can be divided into six subgroups based on these three groups of joint involvement. More than 70% of patients are classified into regular subgroups, indicating that the pattern of synovitis in a patient with RA is stable. When patients who were regularly classified into the first subgroup of patients characterized by no synovitis were removed, more than 60% of patients were still classified into regular subgroups (data not shown), indicating that the stable patterns were observed regardless of activity of RA. As joint destruction was influenced by disease duration, disease activity, and treatment, we analyzed the relative distribution of joint destruction between the three joint groups in a patient with RA. We found that the sixth subgroup of patients, characterized by moderate activity with dominant involvement of large and wrist joints, demonstrated dominant destruction of wrist joints. This suggests that classifying patients with RA into appropriate subgroups would lead to prediction of patterns of joint destruction.

There are reports that evaluating fraction of joints by ultrasonography is a good way to predict future joint damage [11–12]. One study reported that 5 of the 28 joints with MTP2 and MTP5 joints, namely, wrist, MCP2, MCP3, PIP2, and PIP3 joints, are enough for ultrasonography evaluation [12]. Their data seems to be consistent with our results as they selected at least two joints from three different groups into which the 28-joint symptoms were classified. As ultrasonography usually surpasses physical examination in terms of the sensitivity to detect synovitis, it is interesting to analyze whether the assessments of synovitis using ultrasonography show the same pattern of synovitis over the 28 joints in RA.

Our results indicate that RA does not develop synovitis in the 28 joints with the same frequency and that the affected rate of each joint greatly varies from joint to joint. These different distributions of joint synovitis would lead to different distribution of joint destruction. Based on our results, the 28 joints can be categorized into three groups, and it is possible that some fractions of the 28 joints are less informative to assess disease activity than others. It would be interesting to develop a novel simplified joint core set, and analyze the correlation between joint damage and activity score based on this. It would be also interesting to characterize each of RA subsets in more detail.

Materials and Methods

Ethics Statement

Written informed consent to enroll in the database described below was obtained from most of the patients, but for some patients the information regarding the construction of this database was disclosed instead of obtaining written informed consent. Participants who were informed regarding the construction of the database (instead of obtaining written informed consent) were allowed to withdraw from the study if desired.

All data were de-identified and analyzed anonymously. This study was designed in accordance with the Helsinki Declaration. This study including the consent procedure was approved by the ethics committee of Kyoto University Graduate School and Faculty of Medicine.

The KURAMA database

The KURAMA (Kyoto University Rheumatoid Arthritis Management Alliance) database was established in 2011 at Kyoto University to store detailed clinical information and specimens from patients with arthritis and arthropathy. The alliance is composed of rheumatic disease-associated departments in Kyoto University Hospital as well as its allied, integrating previous database and specimen collections in each department and allied. A template for electronic clinical charts developed at Kyoto University Hospital in 2004 to evaluate joint involvements in RA patients was used to obtain joint assessments. Rheumatologists evaluated swelling and tenderness of the 28 joints in patients with RA on each visit and filled in the template. The synovitis information of the 28 joints and data for C-reactive protein and erythrocyte sedimentation rate were extracted from electronic clinical charts [15] and stored in the KURAMA database.

Patients and data of joint assessment

A total of 17,311 joint assessments from 1,314 patients with RA from 2005 to 2011 were obtained in a retrospective manner from the KURAMA database. All of the patients fulfilled ACR revised criteria for RA in 1987 [10] or ACR and EULAR classification criteria for RA in 2010 [16–17].

Analysis of affected frequencies in the 28 joints

RA patients were subdivided depending on whether their data were available in 2011 or not, and the affected frequency in each of the 28 joints was calculated. We compared the order of the affected frequency in the 28 joints between the two patient sets with Spearman's rank-sum coefficient. We separately analyzed the affected rates of joints for swelling and tenderness. When multiple joint assessments in different visits were available in the same patient with RA, we randomly selected one of the assessments as representative in the patient. We compared frequencies between tenderness and swellings for the 28 joints with Spearman's rank-sum coefficient.

Clustering of patients with RA

Clustering analyses were performed by Ward method, using randomly-selected 5,383 evaluations of the 28 joints from 1,314 patients with RA. These evaluations did not contain more than six assessments from each patient to avoid excess influence of particular patients. Affected rates were calculated for the three groups of joints (namely PIP joints, MCP joints and large and wrist joints) in this clustering analysis. For example, when a patient showed tenderness and swelling for all PIP joints, the affected rate of PIP joints in the patient is 2. When a patient showed tenderness for four MCP joints, the affected rate of MCP joints is 0.4.

RA patients were regarded as belonging to a particular group when more than 60% of evaluations belonging to the same patients with four or five evaluations were classified into the same group.

Analysis between RA subgroups and joint destruction

Joint destruction of hand joints in 246 patients with RA was evaluated by modified Sharp score by a trained rheumatologist who was not informed of the patients' characteristics (KM). Joint destruction rates were defined for the three groups of joints as a sum of scores divided by the full score in the joints group. For example, when a patient shows 50 as a sum of scores in the large and wrist group, the patient's joint destruction rate for the group is 0.463 (50/108).

Correlation of the 28 joints and statistical analysis

Correlations of joint symptoms among the 28 joints were estimated separately for tenderness and swelling. We randomly obtained one assessment of the 28 joints in each patient as a representative of the patient's joint assessments for maximization of the power. Kappa coefficient was used to analyze coincidence of joint symptoms in each pair of the 28 joints. Eigen vectors obtained in principal component analysis were used to analyze the deviation of joint symptoms. We resampled joint assessments for each patient and created four other sets of joint assessments. The same correlation analyses were performed using the four resampled assessments to confirm the correlation shown in the first assessment set. Right dominance of the synovitis and joint destruction was analyzed by binomial test. Dominant destruction of joints was evaluated by paired-t test. Statistical analysis was performed by R software or SPSS (ver18).

Supporting Information

Figure S1 Distribution of joint evaluation counts and patients across different years. A) Distribution of number of RA patients according to numbers of 28-joint assessments. B) Distribution of number of patients with RA whose joint assessment data were available from 2005 to 2011 in the KURAMA database. (TIF)

Figure S2 Good correlations between joint involvement rates in different sets of RA patients. Rates of joint involvement for A) swelling and B) tenderness were compared between the two different sets of RA patients. X and Y axes represent rates in the first set of RA patients in 2011 and those in the second set in 2005 to 2010, respectively.
(TIF)

Figure S3 Three groups of joints regardless of different sets of RA patients. Analysis using one of four resampled assessments in one of the two sets of RA patients is shown as a representative. The 1st and 2nd components of eigen vectors of the joint symptoms are plotted, using principal component analysis of the 28 joint involvement for tenderness (A) and swelling (C) or using that of the 20 joint involvement other than large and wrist joints for tenderness (B) and swelling (D). Green: large and wrist joints. Red: MCP joints. Blue: PIP joints.
(TIF)

Figure S4 Three groups of joints regardless of different evaluators. Analysis using one of five resampled assessments by one of the two groups of medical doctors is shown as a representative. The 1st and 2nd components of eigen vectors of the joint symptoms are plotted, using principal component analysis of the 28 joint involvement for tenderness (A) and swelling (C) or using that of the 20 joint involvement other than large and wrist joints for tenderness (B) and swelling (D). Green: large and wrist joints. Red: MCP joints. Blue: PIP joints.
(TIF)

Figure S5 Dominant destruction of large and wrist joints in the sixth subgroup of patients with RA. Box plots indicating the joint destruction rates in the three joint groups in subjects belonging to the sixth subgroup.
(TIF)

Figure S6 Destruction of large and wrist joints among the six subgroups of RA. Differences in destruction rates were plotted for each subject in the six subgroups. The difference was defined as: A) destruction rate of group of large and wrist joints – destruction rate of MCP joints and B) destruction rate of group of large and wrist joints – destruction rate of PIP joints.
(TIF)

Table S1 Rate of joint involvement for 28 joints in RA.
(DOC)

Table S2 Right-dominant joint destruction in RA. Patients who showed unilateral higher or lower scores in each element were analyzed.
(DOC)

Table S3 Mean affected rates of the three joint groups in the six subgroups of patients with RA.
(DOC)

Acknowledgments

We would like to thank to Mr. Wataru Yamamoto at Kurashiki Kosai Hospital for his excellent support to establish and maintain the KURAMA database. We also thank Drs Hisashi Yamanaka, Katsunori Ikari, and Ayako Nakajima at Institute of Rheumatology, Tokyo Women's Medical University for their kind instruction and advice for management of rheumatic diseases database.

Author Contributions

Evaluation of joint X-rays: KM. Conceived and designed the experiments: CT MH KO RY FM HI TF TM. Analyzed the data: CT. Contributed reagents/materials/analysis tools: CT MH KO RN KM N. Yamakawa H. Yoshifuji N. Yukawa DK TU H. Yoshitomi MF HI TF TM KY. Wrote the paper: CT.

References

1. Firestein GS (2003) Evolving concepts of rheumatoid arthritis. Nature 423: 356–361.
2. Drossaers-Bakker KW, de Buck M, van Zeben D, Zwinderman AH, Breedveld FC, et al. (1999) Long-term course and outcome of functional capacity in rheumatoid arthritis: the effect of disease activity and radiologic damage over time. Arthritis and Rheumatism 42: 1854–1860.
3. Smolen JS, Van Der Heijde DM, St Clair EW, Emery P, Bathon JM, et al. (2006) Predictors of joint damage in patients with early rheumatoid arthritis treated with high-dose methotrexate with or without concomitant infliximab: results from the ASPIRE trial. Arthritis and Rheumatism 54: 702–710.
4. Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, et al. (1993) The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. The Committee on Outcome Measures in Rheumatoid Arthritis Clinical Trials. Arthritis and Rheumatism 36: 729–740.
5. van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, et al. (1990) Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. Annals of the Rheumatic Diseases 49: 916–920.
6. van der Heijde DM, van't Hof MA, van Riel PL, van Leeuwen MA, van Rijswijk MH, et al. (1992) Validity of single variables and composite indices for measuring disease activity in rheumatoid arthritis. Annals of the Rheumatic Diseases 51: 177–181.
7. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, et al. (2003) A simplified disease activity index for rheumatoid arthritis for use in clinical practice. Rheumatology 42: 244–257.
8. Aletaha D, Smolen JS (2007) The Simplified Disease Activity Index (SDAI) and Clinical Disease Activity Index (CDAI) to monitor patients in standard clinical care. Best Pract Res Clin Rheumatol 21: 663–675.
9. Salaffi F, Cimmino MA, Leardini G, Gasparini S, Grassi W (2009) Disease activity assessment of rheumatoid arthritis in daily practice: validity, internal consistency, reliability and congruency of the Disease Activity Score including 28 joints (DAS28) compared with the Clinical Disease Activity Index (CDAI). Clinical and Experimental Rheumatology 27: 552–559.
10. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, et al. (1988) The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 31: 315–324.
11. Scheel AK, Hermann KG, Kahler E, Pasewaldt D, Fritz J, et al. (2005) A novel ultrasonographic synovitis scoring system suitable for analyzing finger joint inflammation in rheumatoid arthritis. Arthritis and Rheumatism 52: 733–743.
12. Backhaus M, Ohrndorf S, Kellner H, Strunk J, Backhaus TM, et al. (2009) Evaluation of a novel 7-joint ultrasound score in daily rheumatologic practice: a pilot project. Arthritis and Rheumatism 61: 1194–1201.
13. van der Heijde D (2000) How to read radiographs according to the Sharp/van der Heijde method. Journal of Rheumatology 27: 261–263.
14. Machold KP, Stamm TA, Eberl GJ, Nell VK, Dunky A, et al. (2002) Very recent onset arthritis – clinical, laboratory, and radiological findings during the first year of disease. Journal of Rheumatology 29: 2278–2287.
15. Yamamoto K, Yamanaka K, Hatano E, Sumi E, Ishii T, et al. (2012) An eClinical trial system for cancer that integrates with clinical pathways and electronic medical records. Clin Trials 9: 408–417.
16. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, et al. (2010) 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis and Rheumatism 62: 2569–2581.
17. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, et al. (2010) 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Annals of the Rheumatic Diseases 69: 1580–1588.