

ORIGINAL ARTICLE

Common variants on 14q32 and 13q12 are associated with DLBCL susceptibility

Vinod Kumar^{1,2}, Keitaro Matsuo³, Atsushi Takahashi², Naoya Hosono², Tatsuhiko Tsunoda², Naoyuki Kamatani², Sun-Young Kong⁴, Hidewaki Nakagawa², Ri Cui¹, Chizu Tanikawa¹, Masao Seto⁵, Yasuo Morishima⁶, Michiaki Kubo², Yusuke Nakamura^{1,2} and Koichi Matsuda¹

Diffuse large B-cell lymphoma (DLBCL) is one of the most aggressive cancers of B-lymphocytes. To investigate genetic susceptibility factors for DLBCL, we performed single-nucleotide polymorphism based genome-wide association study (GWAS) in a total of 399 DLBCL cases and 4243 controls of Japanese population. By following two-stage GWAS approach and an independent replication study, we identified disease susceptibility locus within intron 3 of the *CDC42BPB* gene on 14q32 (rs751837; $P=3.30 \times 10^{-7}$ and odds ratio (OR) of 3.5), a region of frequent chromosomal translocations in lymphoma, and variant on 13q12 (rs7097; $P=6.57 \times 10^{-6}$ and OR of 1.43) which harbors the notch signaling mediator, *LNX2* gene. Our findings would contribute to the understanding of DLBCL risk and also may lead to the elucidation of its molecular pathogenesis.

Journal of Human Genetics (2011) 56, 436–439; doi:10.1038/jhg.2011.35; published online 7 April 2011

Keywords: association study; DLBCL; lymphoma; 13q12; 14q32

INTRODUCTION

Diffuse large B-cell lymphoma (DLBCL), the major subtype of non-Hodgkin's B-cell lymphoma, is accounting for >30% of all lymphoid malignancies worldwide¹ and >43% in Japan.² Although the combination of cyclophosphamide, doxorubicin, vincristine, prednisone and rituximab, has extended the overall survival among DLBCL patients,³ a substantial portion of patients still suffer from persistent or recurrent diseases. To further improve the clinical outcome of DLBCL, elucidation of its underlying pathogenesis is essential.

Along with the established environmental risk factors such as older age, infectious diseases and congenital immunodeficiency, genetic factors also have a major role.^{4–6} According to this notion, nearly 10-fold relative risk for DLBCL was reported among first-degree relatives of DLBCL patients.⁷ Thus, it is conceivable that the pathogenesis of this aggressive disease is largely determined by complex interaction between genetic and environmental factors. To clarify the genetic susceptibility factors for DLBCL, we conducted single-nucleotide polymorphism (SNP) based genome-wide association study (GWAS) among clinically and histopathologically confirmed DLBCL patients.

MATERIALS AND METHODS

Study participants

In this study, we conducted two-stage GWAS and a subsequent replication analysis using a total of 399 DLBCL cases and 4243 control subjects. The

demographic details of study participants are summarized in Supplementary Table 1. In the first stage of GWAS, 934 Japanese control DNA samples were obtained from Osaka-Midosuji Rotary Club, Osaka, Japan. DLBCL cases in the first and the second stage and control subjects in the second stage were obtained from BioBank Japan.⁸ Histopathologically confirmed 106 DLBCL cases and age-gender-matched 400 healthy controls were selected from Aichi Cancer Center, Japan for a replication study.⁹ All the participants provided written informed consent. This research project was approved by the ethics committees at each institute.

SNP genotyping

In all, 74 DLBCL cases and 934 healthy controls in the first stage and 2909 controls in the second stage were genotyped using Illumina HumanHap550v3 Genotyping BeadChip (Illumina, San Diego, CA, USA). Standard SNP quality control filters (P -value of Hardy–Weinberg equilibrium test of $>1.0 \times 10^{-6}$ for controls, minor allele frequency of >0.01 , genotyping completeness of $\geq 99\%$) were applied to finally obtain 444 361 SNPs for the association analysis. Case samples in the second stage and case–control samples in the replication analysis were genotyped by the multiplex PCR-based Invader assay (Third Wave Technologies, Inc., Madison, WI, USA).¹⁰

Statistical analysis

The association was tested by logistic regression analysis after adjusting for age and gender in the first stage of GWAS. At the second stage and the replication analysis, statistical significance of the association with each SNP was assessed

¹Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, the University of Tokyo, Tokyo, Japan; ²Center for Genomic Medicine, RIKEN, Kanagawa, Japan; ³Division of Epidemiology and Prevention, Aichi Cancer Center Central Hospital, Nagoya, Japan; ⁴Department of Laboratory Medicine, Center for Clinical Services, Research Institute and Hospital, National Cancer Center, Gyeonggi-do, Korea; ⁵Division of Molecular Medicine, Aichi Cancer Center Central Hospital, Nagoya, Japan and ⁶Department of Hematology and Cell Therapy, Aichi Cancer Center Central Hospital, Nagoya, Japan

Correspondence: Dr K Matsuda, Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, 4-6-1, Shirokanedai, Minato, Tokyo 1088639, Japan.

E-mail: koichima@ims.u-tokyo.ac.jp

Received 6 December 2010; revised 28 January 2011; accepted 8 March 2011; published online 7 April 2011

using 1 d.f. Cochran–Armitage trend test. Odds ratios (OR) and confidence intervals (CIs) were calculated using the major allele as a reference allele. The meta-analysis was conducted using the Mantel–Haenszel method. Heterogeneity among studies was examined by the Breslow–Day test. Software and web tools used in this study are summarized in Supplementary data.

Candidate gene analysis

We searched pubmed for association studies of lymphoma, and obtained the list of 182 candidate genes those were previously analyzed by several researchers among different population by candidate gene approach. We selected tag SNPs located on these genes from Illumina 550k platform (Illumina; Supplementary Table 3). The association between these tag SNPs and DLBCL was tested in the first stage 74 DLBCL cases and 934 healthy controls by logistic regression analysis on adjustment for age and gender.

Software and web tools

For general statistical analyses, we used R statistical environment version 2.6.1 (cran.r-project.org) or plink-1.05 (pngu.mgh.harvard.edu/~purcell/plink).¹¹ The Haploview software was used to draw linkage disequilibrium maps.¹² Primer3-web v0.3.0 (http://frodo.wi.mit.edu) web tool was used to design primers. The r^2 between SNPs were annotated using the WGAViewer software.¹³ In-silico functional annotation of SNPs was performed using FastSNP web tool (http://fastsnp.ibms.sinica.edu.tw/fastSNP2/pages/input_CandidateGeneSearch.jsp).¹⁴

RESULTS

First-stage GWAS

In the first stage, we genotyped 74 DLBCL cases and 934 healthy control samples at 550 000 SNP loci. The association between each SNP and DLBCL was assessed by logistic regression analysis after adjusting for age and gender. We plotted observed logistic P -values against expected logistic P -values by quantile–quantile plot and confirmed that there is no or least population stratification ($\lambda=1.03$, Supplementary Figure 1). To identify disease susceptibility loci, we further investigated top ranked 500 SNPs ($P \leq 1 \times 10^{-3}$, Supplementary Figure 2) in additional cases and controls.

Second-stage GWAS

The analysis of the second stage cohort which consisted of 219 DLBCL cases and 2909 controls at 500 top SNPs revealed significant association (Armitage $P < 0.05$) at 15 SNPs (Supplementary Table 2). The Mantel–Haenszel test was performed at these 15 SNPs to assess the overall degree of association by combining first stage and second stage

datasets. Notably, SNP rs4551233 and SNP rs4443228 showed strong association with DLBCL. Both SNPs achieved the Mantel–Haenszel P of 7.06×10^{-7} and 7.23×10^{-7} ; OR of 1.57 (95% CI 1.32–1.88) and 2.43 (95% CI 1.7–3.45), respectively. We also observed improved associations at SNPs rs11222532, rs17811655, rs1381795 and rs751837. However, no SNPs exceeded a genome-wide significant threshold ($P < 1.13 \times 10^{-7}$, after Bonferroni correction).

Replication analysis using Aichi cancer center DLBCL cohort

In all, 15 candidate SNPs from our two-stage GWAS were subsequently genotyped in age and gender matched independent DLBCL replication cohort, consisting of 106 cases and 400 controls from Aichi Cancer Center. Genotyping of this cohort revealed significant association (Table 1) at rs7097 ($P=4.89 \times 10^{-2}$, OR of 1.37; 95% CI 1.01–1.87) that is located on chromosome 13q12–q13 (Figure 1a). In addition, we observed nominal association at rs751837 ($P=6.19 \times 10^{-2}$) that is located within 120 kb linkage disequilibrium block on chromosome 14q32.32 (Figure 1b). The meta-analysis revealed consistent association at rs7097 ($P=6.57 \times 10^{-6}$ and OR of 1.43; 95% CI=1.23–1.67) and rs751837 ($P=3.3 \times 10^{-7}$ and OR of 3.5; 95% CI=2.12–5.88) with the same direction of effect (Supplementary Figure 3).

DISCUSSION

Our study identified two genetic variants to be consistently associated with increased risk of DLBCL susceptibility. SNP rs7097 is located at the common region encompassing the 3-prime un-translated region of *POLR1D* and the promoter region of *LNX2* (ligand of numb protein $\times 2$; Figure 1a). *POLR1D* encodes a 16 kDA RNA polymerase I polypeptide D. By considering the role of POLR1D in transcription, it is hard to explain the impact of its polymorphisms on DLBCL pathogenesis. On the other hand, *LNX2* encodes a PDZ domain containing ring finger 1 protein which may function as an E3 ubiquitin ligase.¹⁵ Higher levels of LNX family member was shown to cause ubiquitin–proteasomal degradation of Numb protein which in turn enhance the notch signaling.¹⁵ As activated notch signaling is implicated in hematological cancers¹⁶ such as B-cell lymphoma,¹⁷ SNP rs7097 might have functional consequences for DLBCL. Although, on the basis of *in silico* prediction, allele G at SNP rs7097 can create a binding site for GATA-2 transcription factor (Supplementary Figure 4A), identifying a specific causative variant at this locus is necessary to understand the impact of this locus for DLBCL.

Table 1 Meta-analysis at two candidates by combining first stage, second stage and replication stage

rs ID	Gene	Chromosome	Stage	Case				Control				Armitage P	M–H P	OR (95% CI)	P _{hetero} ^a	
				GG	AG	AA	MAF	GG	AG	AA	MAF					
rs7097	LNX2	13	First stage	7	32	35	0.311	189	486	259	0.463	2.73×10^{-4}				
			Second stage	26	116	76	0.385	584	1453	838	0.456	3.88×10^{-3}				
			Replication	21	44	40	0.410	97	196	107	0.488	4.89×10^{-2}				
			Combined	54	192	151	0.378	870	2135	1204	0.460	6.57×10^{-6}	3.09×10^{-6}	1.437 (1.23–1.67)	0.73	
				CC	CT	TT	MAF	CC	CT	TT	MAF					
rs751837	CDC42BPB	14	First stage	0	7	67	0.047	0	14	920	0.007	3.93×10^{-6}				
			Second stage	0	8	200	0.019	0	43	2813	0.008	1.09×10^{-2}				
			Replication	1	3	101	0.024	0	6	393	0.008	6.19×10^{-2}				
			Combined	1	18	368	0.026	0	63	4126	0.008	3.30×10^{-7}	1.35×10^{-6}	3.508 (2.127–5.88)	0.86	

Abbreviations: CI, confidence interval; M–H, Mantel–Haenszel; OR, odds ratio.
^aP determined by Breslow–Day test for heterogeneity across three stages.

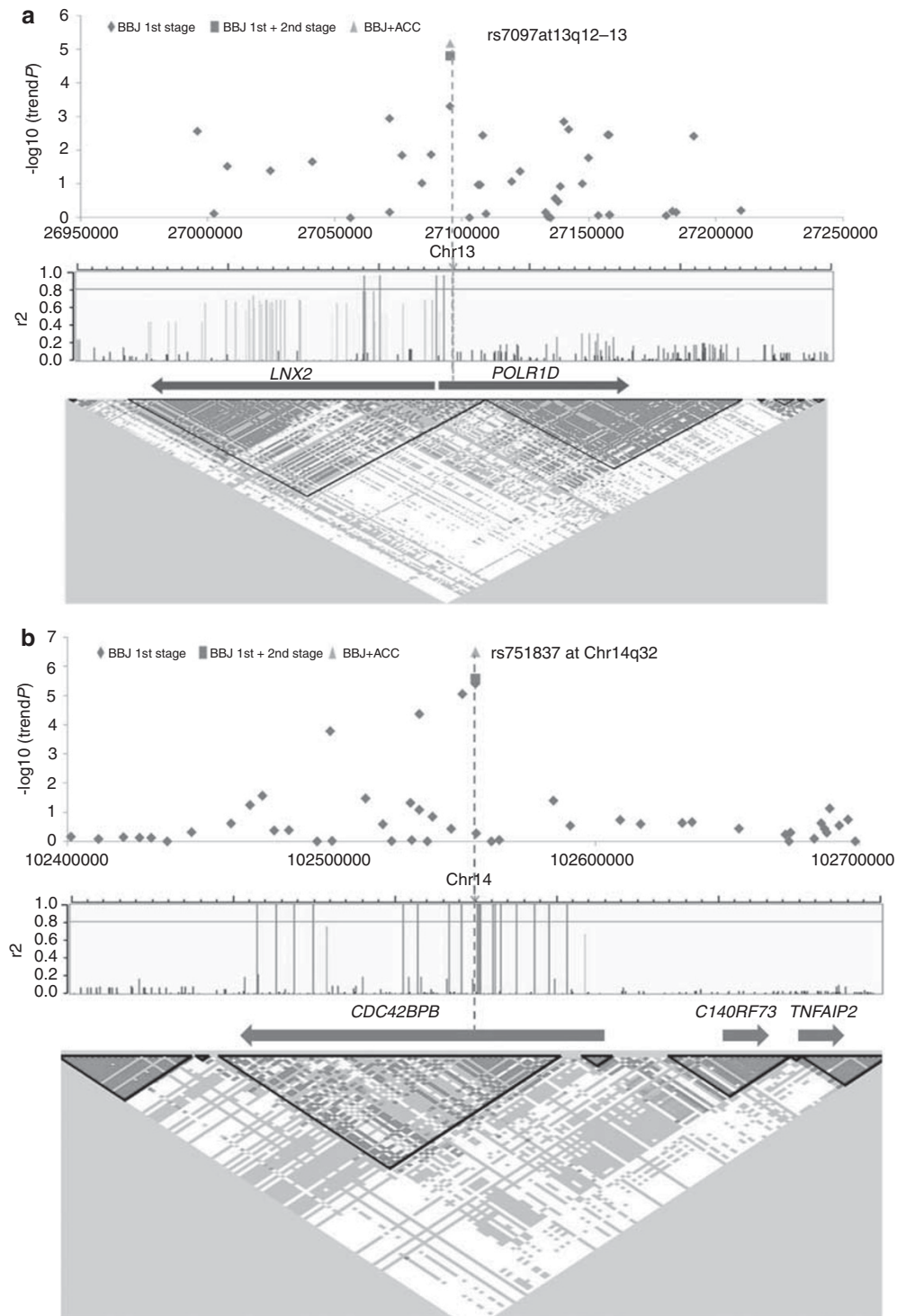


Figure 1 (a) Regional association plot on 13q12–13. The upper panel displays the distribution of $-\log_{10} P$ -values (Cochran–Armitage trend test) according to their physical positions on 13q12–13. The dotted vertical line indicates the physical position of rs7097 SNP and its combined P -value. The second panel shows the r^2 values for proxies of rs7097, which are based on the HapMap JPT population. Annotated genes are shown at the top of linkage disequilibrium block. (b) Regional association plot on 14q32. The upper panel displays the distribution of $-\log_{10} P$ -values (Cochran–Armitage trend test) according to their physical position on 14q32. The dotted vertical line indicates the physical position of rs751837 SNP and its combined P -value. The second panel shows the r^2 values for proxies of rs751837, which are based on the HapMap JPT population. Annotated genes are shown at the top of linkage disequilibrium block. BBJ, biobank Japan; ACC, Aichi Cancer Center. A full color version of this figure is available at the *Journal of Human Genetics* journal online.

The second candidate SNP rs751837 is located within the intron 3 of *CDC42BPB* (also known as MRCK β (myotonic dystrophy kinase-related Cdc42-binding kinase β) Figure 1b). According to *in silico*

prediction, allele T at rs751837 shows preferential binding of the CdxA transcription factor (Supplementary Figure 4b) implying possible role of this variant on the transcriptional regulation of *CDC42BPB*.

However, further characterization of this region is required to delineate the effect of other variants on this locus. CDC42BPB is a member of the multidomained serine/threonine protein kinase family. MRCK isoforms affects cellular structures through modulating Cdc42 actions during cytoskeletal reorganization,¹⁸ and are shown to interact with potent tumor-promoting agent phorbol ester and diacylglycerol, albeit with weaker affinity.¹⁹ Thus, CDC42BPB might have a role in carcinogenesis by transducing diacylglycerol signals similar to protein kinase C or RasGRPs. In addition, CDC42BPB is located on chromosome 14q32 region, which is inherently susceptible to frequent chromosomal translocations, especially in lymphoma. It is reported that the presence of 14q32 translocation among DLBCL is associated with higher frequencies of other chromosomal aberrations.²⁰ Thus, further analysis about correlation between rs751837 and 14q32 translocation might shed light on the molecular mechanism underlying lymphoma specific chromosomal translocation.

The SNPs identified in this study, however failed to achieve conservative Bonferroni threshold ($P < 5 \times 10^{-8}$) for GWAS, which could be because of the small sample size analyzed in this study and/or the clinical heterogeneity of DLBCL. In this regard, it is noteworthy that a recent GWAS among Caucasians identified 6p21.33 locus as a susceptibility region to follicular lymphoma, however, the same study failed to identify genetic susceptibility factors to DLBCL with a relatively larger sample size (783 cases and 3377 controls).²¹ In contrast, we found two DLBCL susceptibility loci those seem to be functionally related to lymphoma pathogenesis. This is probably because of the genetic homogeneity among Japanese population or the difference in the genetic background between Asian and Caucasian.²²

We also attempted to determine the association between DLBCL and candidate genes analyzed by several researchers from different population. The analysis revealed nominal association ($P < 0.05$) at SNPs located on 52 candidate genes (Supplementary Table 3), which are mainly implicated in immune response, apoptosis and cytokine signaling. These finding suggested that many genetic variations would contribute to the pathogenesis of DLBCL.

As DLBCL exhibit clinical, histological, immunological and molecular heterogeneity,^{23–25} more detailed DLBCL classification (for example, germinal center DLBCL or non-germinal center DLBCL or activated B-cell-like) may be necessary to identify subtype specific genetic variants both from candidate gene analysis as well as from GWAS. In this regard, a global collaborative effort might be required to obtain sufficient samples to undertake a meta-analysis to identify stronger genetic susceptibility factors for subtypes of DLBCL. Nevertheless, our findings support the role of genetic component in DLBCL susceptibility and may contribute toward understanding of DLBCL pathogenesis.

ACKNOWLEDGEMENTS

This work was conducted as a part of the BioBank Japan Project that was supported by the Ministry of Education, Culture, Sports, Sciences and Technology of the Japanese government. We thank all the patients who donated their DNA for this study and the Rotary Club of Osaka-Midosuji District 2660 Rotary International in Japan for providing DNA samples of healthy controls.

- 1 Coiffier, B. Immunotherapy: the new standard in aggressive non-Hodgkin's lymphoma in the elderly. *Semin. Oncol.* **30**, 21–27 (2003).
- 2 Izumo, T., Maseki, N., Mori, S. & Tsuchiya, E. Practical utility of the revised European-American classification of lymphoid neoplasms for Japanese non-Hodgkin's lymphomas. *Jpn. J. Cancer Res.* **91**, 351–360 (2000).
- 3 Coiffier, B., Lepage, E., Briere, J., Herbrecht, R., Tilly, H., Bouabdallah, R. et al. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. *N. Engl. J. Med.* **346**, 235–242 (2002).
- 4 Altieri, A., Bermejo, J. L. & Hemminki, K. Familial risk for non-Hodgkin lymphoma and other lymphoproliferative malignancies by histopathologic subtype: the Swedish Family-Cancer Database. *Blood* **106**, 668–672 (2005).
- 5 Lech-Maranda, E., Baseggio, L., Bienvenu, J., Charlot, C., Berger, F., Rigal, D. et al. Interleukin-10 gene promoter polymorphisms influence the clinical outcome of diffuse large B-cell lymphoma. *Blood* **103**, 3529–3534 (2004).
- 6 Habermann, T. M., Wang, S. S., Maurer, M. J., Morton, L. M., Lynch, C. F., Ansell, S. M. et al. Host immune gene polymorphisms in combination with clinical and demographic factors predict late survival in diffuse large B-cell lymphoma patients in the pre-rituximab era. *Blood* **112**, 2694–2702 (2008).
- 7 Goldin, L. R., Björkholm, M., Kristinsson, S. Y., Turesson, I. & Landgren, O. Highly increased familial risks for specific lymphoma subtypes. *Br. J. Haematol.* **146**, 91–94 (2009).
- 8 Nakamura, Y. The BioBank Japan project. *Clin. Adv. Hematol. Oncol.* **5**, 696–697 (2007).
- 9 Matsuo, K., Hamajima, N., Suzuki, R., Ogura, M., Kagami, Y., Taji, H. et al. Methylene-tetrahydrofolate reductase gene (MTHFR) polymorphisms and reduced risk of malignant lymphoma. *Am. J. Hematol.* **77**, 351–357 (2004).
- 10 Ohnishi, Y., Tanaka, T., Ozaki, K., Yamada, R., Suzuki, H. & Nakamura, Y. A high-throughput SNP typing system for genome-wide association studies. *J. Hum. Genet.* **46**, 471–477 (2001).
- 11 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
- 12 Barrett, J., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265 (2005).
- 13 Ge, D., Zhang, K., Need, A. C., Martin, O., Fellay, J., Urban, T. J. et al. WGAViewer: software for genomic annotation of whole genome association studies. *Genome Res.* **18**, 640–643 (2008).
- 14 Yuan, H. Y., Chiou, J., Tseng, W., Liu, C. H., Liu, C. K., Lin, Y. J. et al. FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization. *Nucleic Acids Res.* **34**, W635–W641 (2006).
- 15 Nie, J., McGill, M. A., Dermer, M., Dho, S. E., Wolting, C. D. & McGlade, C. J. LNX functions as a RING type E3 ubiquitin ligase that targets the cell fate determinant Numb for ubiquitin-dependent degradation. *EMBO J.* **21**, 93–102 (2002).
- 16 Leong, K. G. & Karsan, A. Recent insights into the role of Notch signaling in tumorigenesis. *Blood* **107**, 2223–2233 (2006).
- 17 Chiaramonte, R. Still puzzling Notch signaling in B-cell malignancies. *Leuk. Res.* **30**, 1331–1332 (2006).
- 18 Leung, T., Chen, X. Q., Tan, I., Manser, E. & Lim, L. Myotonic dystrophy kinase-related Cdc42-binding kinase acts as a Cdc42 effector in promoting cytoskeletal reorganization. *Mol. Cell. Biol.* **18**, 130–140 (1998).
- 19 Choi, S. H., Czifra, G., Kedei, N., Lewin, N. E., Lazar, J., Pu, Y. et al. Characterization of the interaction of phorbol esters with the C1 domain of MRCK (myotonic dystrophy kinase-related Cdc42 binding kinase) alpha/beta. *J. Biol. Chem.* **283**, 10543–10549 (2008).
- 20 Cigudosa, J. C., Parsa, N. Z., Louie, D. C., Filippa, D. A., Jhanwar, S. C., Johansson, B. et al. Cytogenetic analysis of 363 consecutively ascertained diffuse large B-cell lymphomas. *Genes Chromosomes Cancer* **25**, 123–133 (1999).
- 21 Skibola, C. F., Bracci, P. M., Halperin, E., Conde, L., Craig, D. W., Agana, L. et al. Genetic variants at 6p21.33 are associated with susceptibility to follicular lymphoma. *Nat. Genet.* **41**, 873–875 (2009).
- 22 Yamaguchi-Kabata, Y., Nakazono, K., Takahashi, A., Saito, S., Hosono, N., Kubo, M. et al. Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. *Am. J. Hum. Genet.* **83**, 445–456 (2008).
- 23 Gurbuxani, S., Anastasi, J. & Hyjek, E. Diffuse large B-cell lymphoma—more than a diffuse collection of large B cells: an entity in search of a meaningful classification. *Arch. Pathol. Lab. Med.* **133**, 1121–1134 (2009).
- 24 Hans, C. P., Weisenburger, D. D., Greiner, T. C., Gascoyne, R. D., Delabie, J., Ott, G. et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood* **103**, 275–282 (2004).
- 25 Lenz, G., Wright, G., Dave, S. S., Xiao, W., Powell, J., Zhao, H. et al. Stromal gene signatures in large-B-cell lymphomas. *N. Engl. J. Med.* **359**, 2313–2323 (2008).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)