

The potential of copy number gains and losses, detected by array-based comparative genomic hybridization, for computational differential diagnosis of B-cell lymphomas and genetic regions involved in lymphomagenesis

Ichiro Takeuchi,¹ Hiroyuki Tagawa,² Akira Tsujikawa,³ Masao Nakagawa,² Miyuki Katayama-Suguro,^{3,2} Ying Guo,^{2,4} and Masao Seto^{2,5}

¹Department of Scientific and Engineering Simulation, Nagoya Institute of Technology, Nagoya, Japan; ²Division of Molecular Medicine, Aichi Cancer Center Research Institute, Nagoya, Japan; ³Division of Information Engineering, Graduate School of Engineering, Mie University, Tsu, Japan; ⁴Department of Pathology, State Key Laboratory of Cancer Biology, Xijing Hospital, Fourth Military Medical University, Shaanxi, P.R.China; ⁵Department of Cancer Genetics, Nagoya University Graduate School of Medicine at Aichi Cancer Center, Nagoya, Japan

Citation: Takeuchi I, Tagawa H, Tsujikawa A, Nakagawa M, Katayama-Suguro M, Guo Y, and Seto M. The potential of copy number gains and losses, detected by array-based comparative genomic hybridization, for computational differential diagnosis of B-cell lymphomas and genetic regions involved in lymphomagenesis. Haematologica 2009. doi: 10.3324/haematol.12986

Description of Supplementary Appendix, Data and Figures

All the following supplementary materials can be downloaded from <http://www-nkn.ics.nitech.ac.jp/~takeuchi/ACGH> with username: *guest* and password: *acghclassifier*.

Refer to related PDF for all the Supplementary Figures cited.

Supplementary Appendix 1

Data analysis procedures. Details of the data analysis procedures including data preprocessing, classifier design, validation and significance test.

Supplementary Data 1

The preprocessed log₂-ratio data for the 46 cases of DLBCL and the 29 cases of MCL.

Supplementary Data 2

The preprocessed log₂-ratio data for the 28 cases of ABC DLBCL and the cases of 18 GCB subtype.

Supplementary Data 3

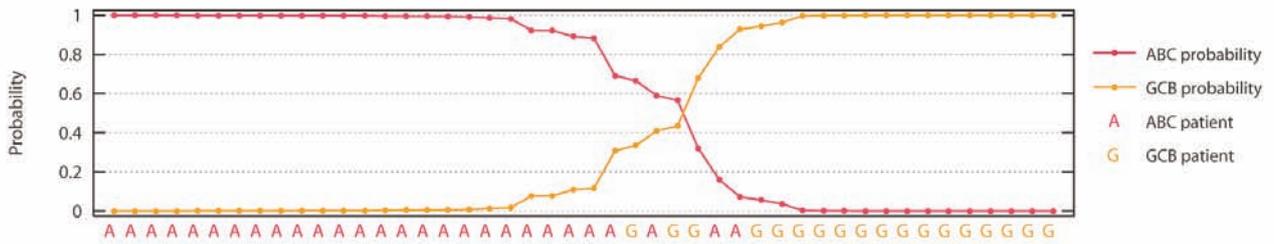
This file contains the preprocessed log₂-ratio signals at 2035 clones for cases of 28 ABC DLBCL and the 18 cases of GCB subtype. A column represents a case and a row represents a clone. Column A: clone ID, column B: clone name, column C: cytogenetic position, column D: known genes contained in the clone, columns E-AF: ABC cases, columns AG-AX: GCB cases, row 1: descriptions, rows 2-2036: clones.

Supplementary Data 4

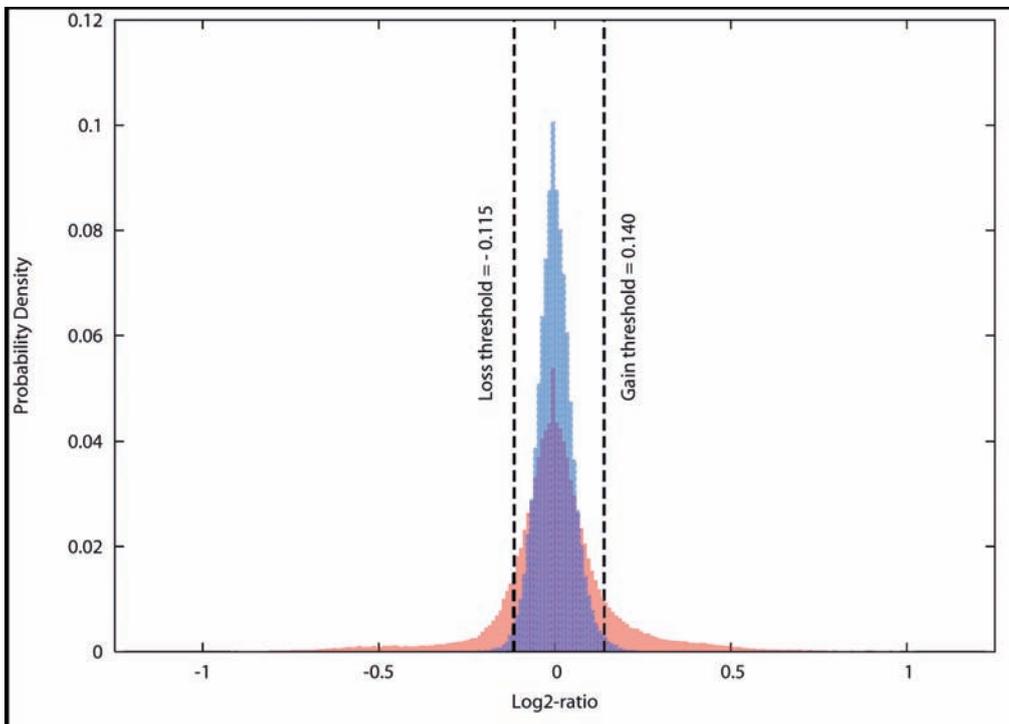
Detailed results of the classification of DLBCL versus MCL.

Supplementary Data 5

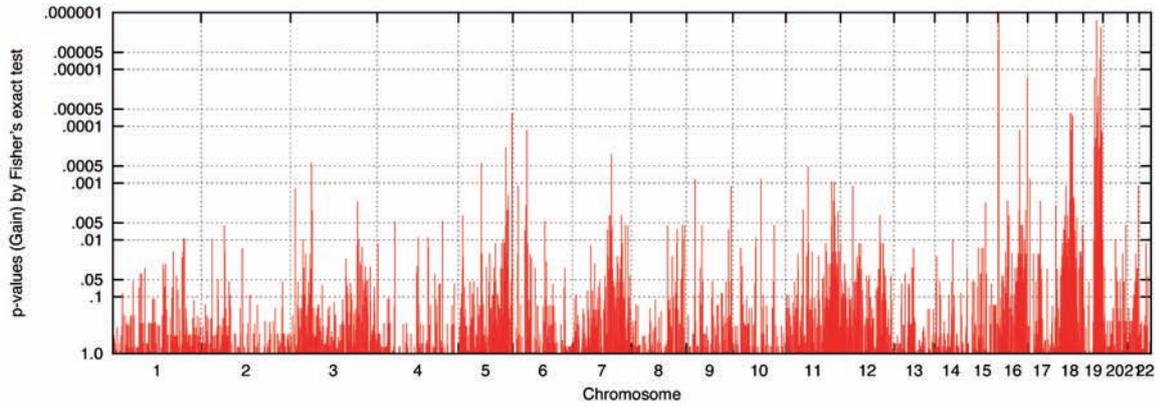
Detailed results of classification of ABC versus GCB subtypes of DLBCL.



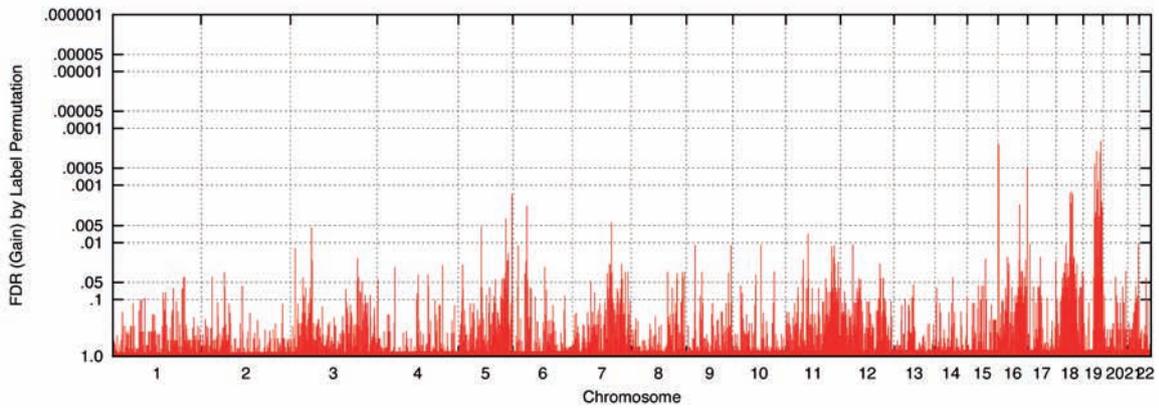
Online Supplementary Figure S1. Confirmation of DLBCL subtype classification. The ABC and GCB subtypes for the analysis were classified using the hierarchical clustering method.¹ The validity of subtype classification was confirmed with an alternative method described by Wright *et al.*¹ The figure shows the probabilities of diagnosing ABC and GCB subtypes using with the gene expression-based classifier. The classification accuracy of the gene expression-based classifier determined by LOOCV was 91.3% (95%CI: 0.832 - 0.994). Two ABC cases were mis classified as GCB and two GCB cases were mis classified as ABC. With an 80% cut-off level, only two cases were mis diagnosed, suggesting that the validity of the ABC and GCB classification used in the our study is strongly supported by an alternative classification method for gene-expression analysis.



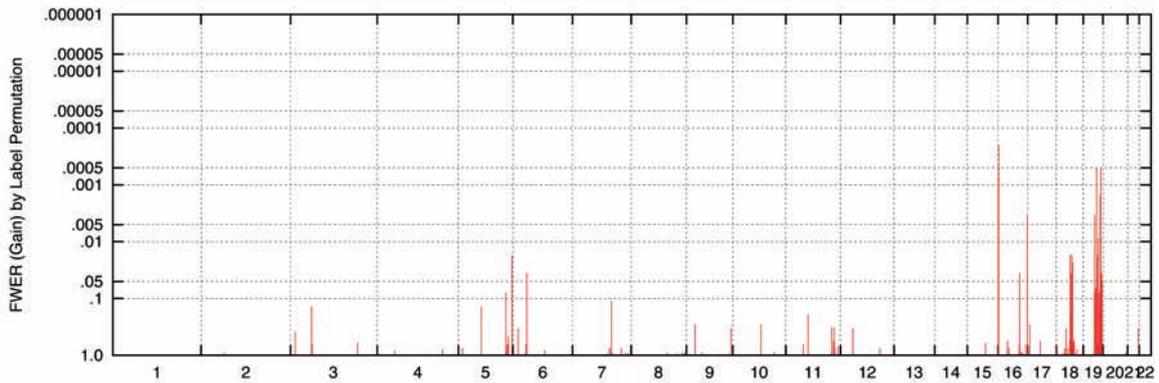
Online Supplementary Figure S2. Empirical distributions of log₂-ratio signals. Empirical distributions of preprocessed log₂-ratio signals for tumor cases (red) and normal cases (blue) as well as the thresholds for copy number gain and loss. The thresholds were set to make the false discovery rate for copy number gains and losses 5%.



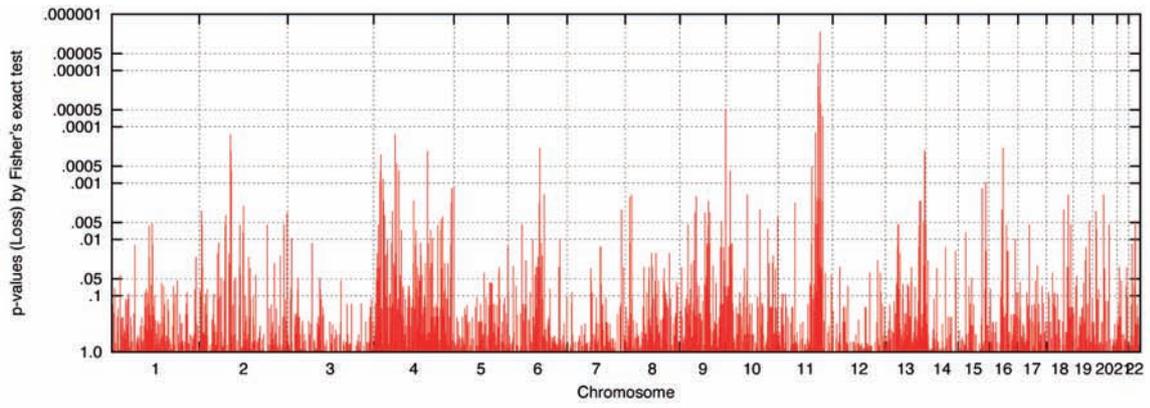
Fisher's exact p for copy number gains in the classification of DLBCL and MCL



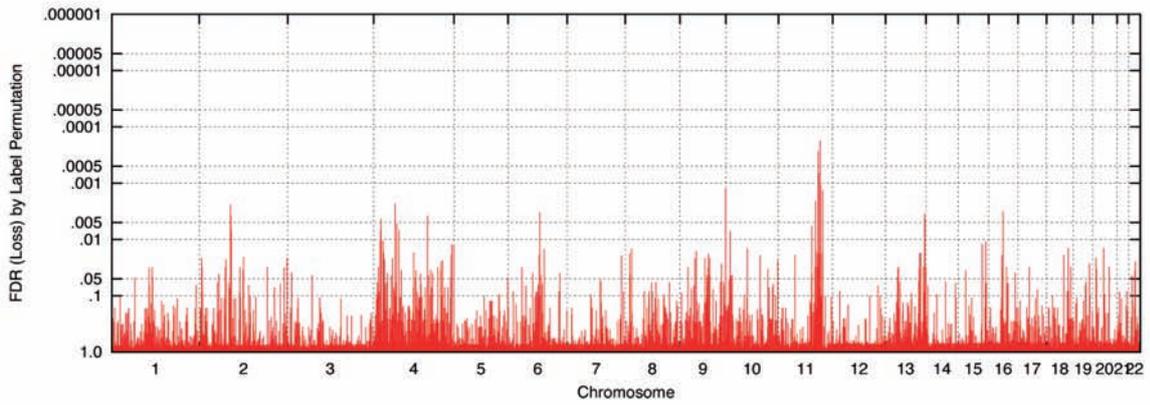
FDR for copy number gains in the classification of DLBCL and MCL



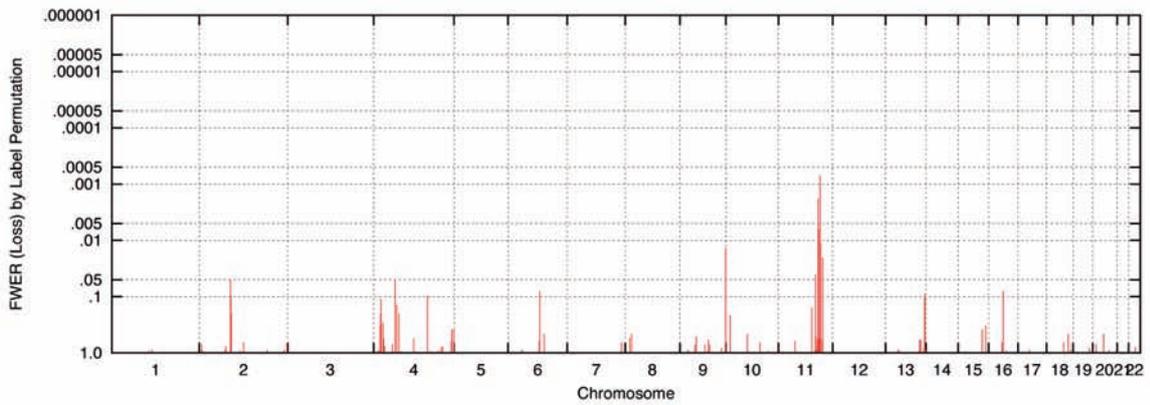
Online Supplementary Figure S3A. Significance of the difference in changes in copy numbers between DLBCL and MCL. (A) p values from Fisher's exact test, false discovery rate (FDR) and family-wise error rate (FWER) for the differences in copy number gains between DLBCL and MCL.



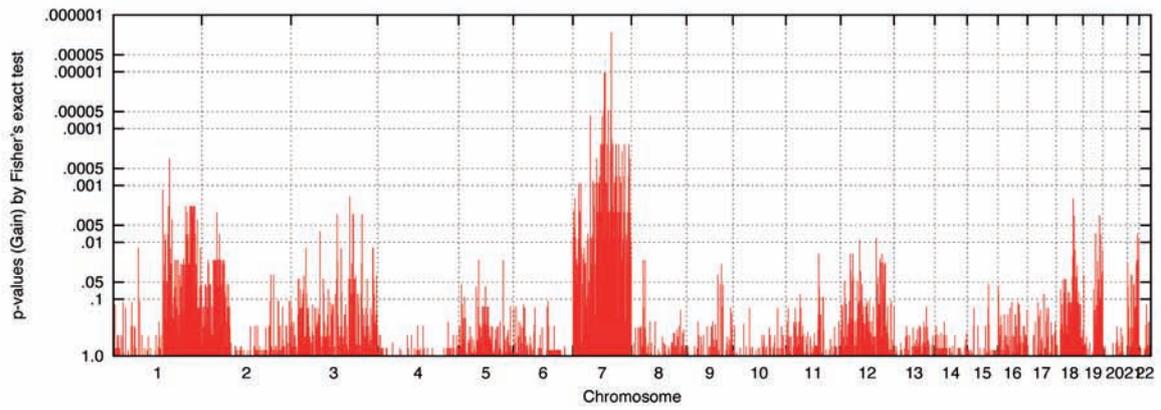
Fisher's exact p for copy number losses in the classification of DLBCL and MCL



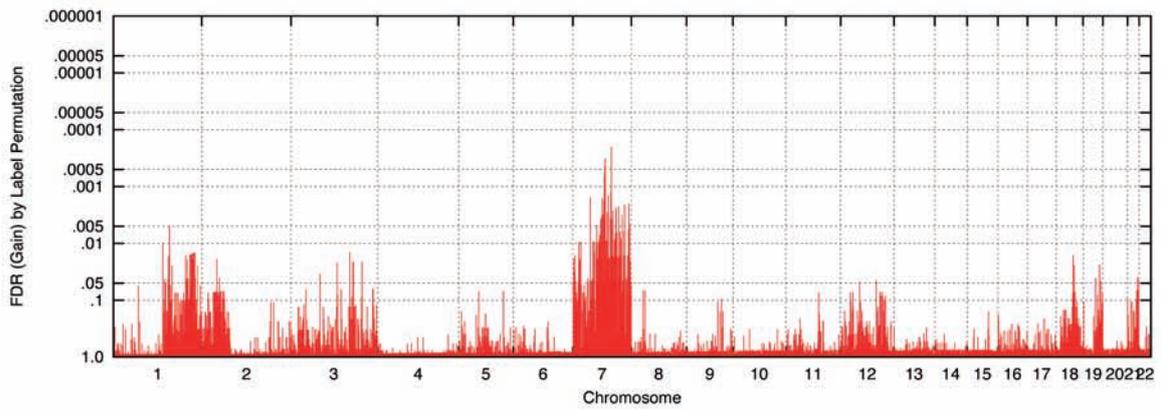
FDR for copy number losses in the classification of DLBCL and MCL



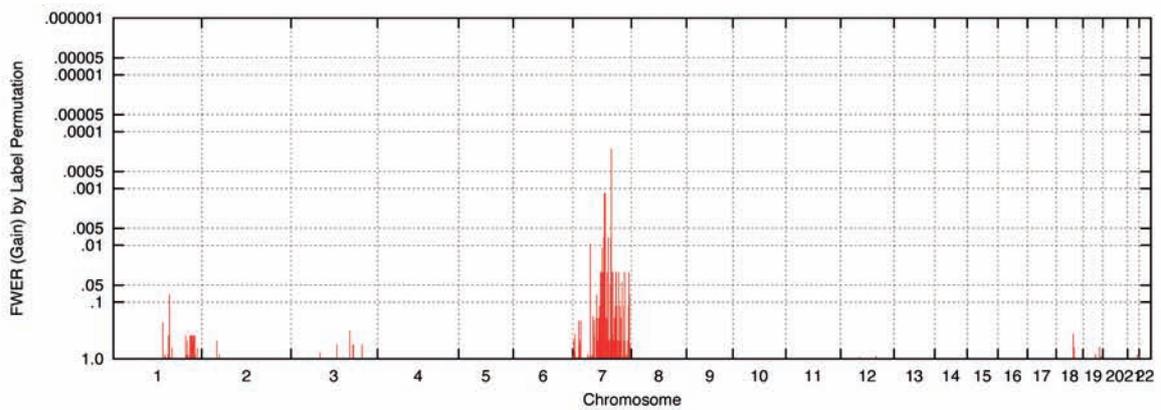
Online Supplementary Figure S3B. Significance of the difference in changes in copy numbers between DLBCL and MCL. (B) p values from Fisher's exact test, FDR and FWER for the differences in copy number losses between DLBCL and MCL.



Fisher's exact p for copy number gains in the classification of ABC and GCB subtypes of DLBCL

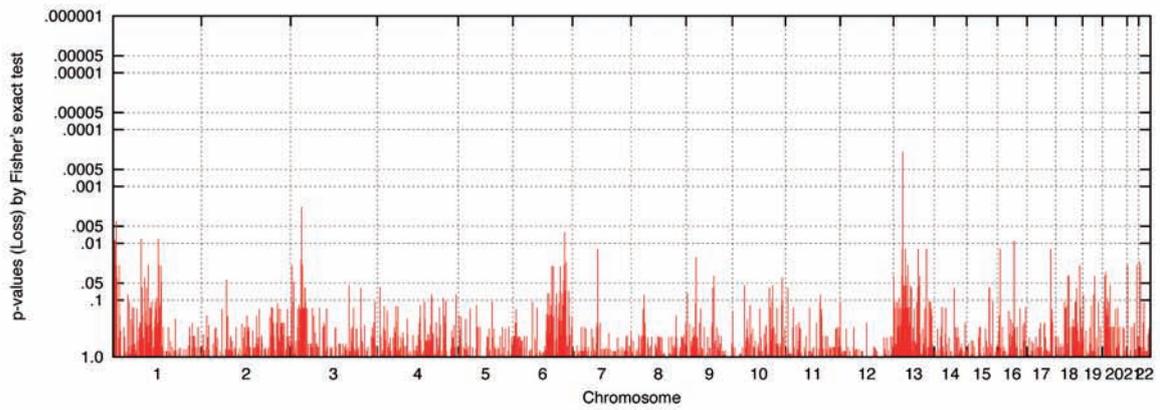


FDR for copy number gains in the classification of ABC and GCB subtypes of DLBCL

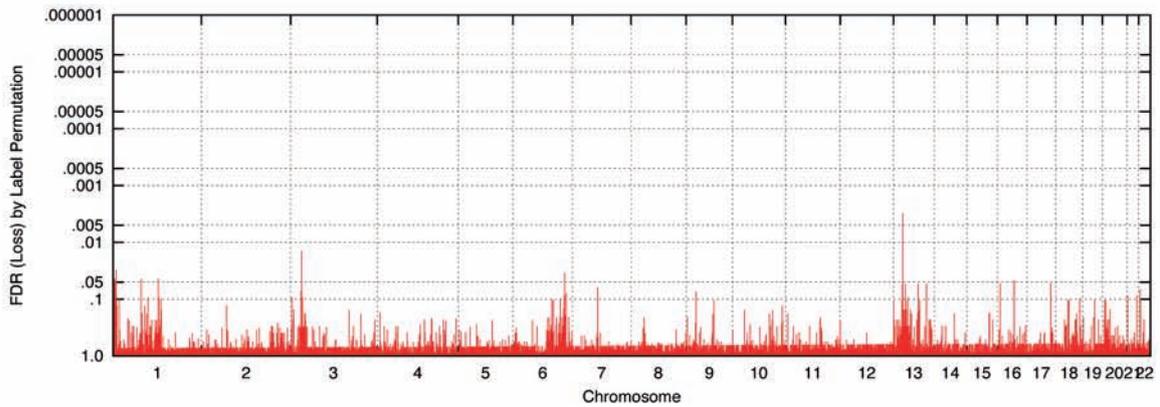


FWER for copy number gains in the classification of ABC and GCB subtypes of DLBCL

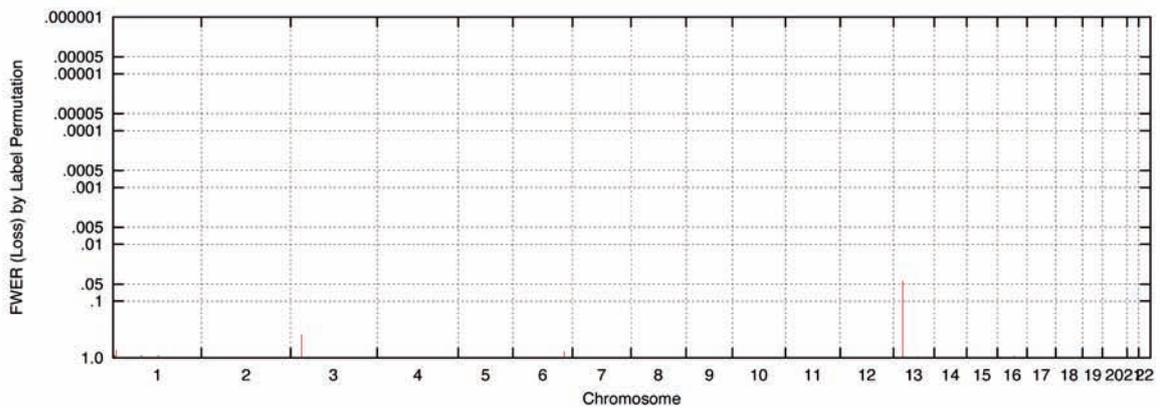
Online Supplementary Figure S4A. Significance of the difference in changes in copy numbers between ABC and GCB subtypes. (A) p values from Fisher's exact test, false discovery rate (FDR) and family-wise error rate (FWER) for the differences in copy number gains between ABC and GCB subtypes of DLBCL.



Fisher's exact p for copy number losses in the classification of ABC and GCB subtypes of DLBCL



FDR for copy number losses in the classification of ABC and GCB subtypes of DLBCL



FWER for copy number losses in the classification of ABC and GCB subtypes of DLBCL

Online Supplementary Figure S4B. Significance of the difference in changes in copy numbers between ABC and GCB subtypes. (B) p values from Fisher's exact test, FDR and FWER for the differences in copy number losses between ABC and GCB subtypes, of DLBCL.
