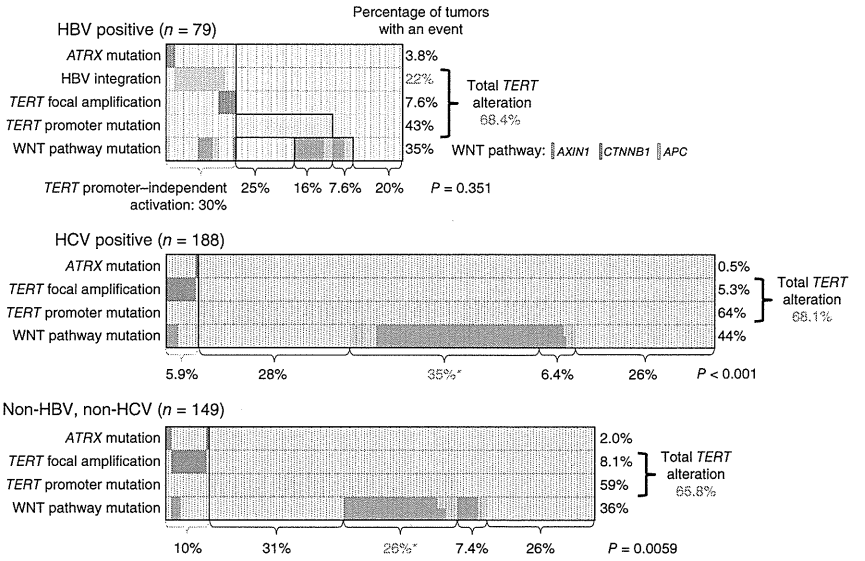

Figure 1 Multiple types of *TERT* alterations in HCC. Mutual exclusivity of HBV genome integration at the *TERT* locus, *TERT* focal amplification and *TERT* promoter mutation in HBV-positive (top), HCV-positive (middle) and non-HBV, non-HCV (bottom) cases. *AXIN1*, *CTNNB1* and *APC* mutations were included as WNT pathway mutations. *TERT* promoter mutation significantly co-occurred with WNT pathway mutation in HBV-negative cases (*$P < 0.001$, χ^2 test). HBV-positive cases without virus capture analysis (41 samples) were excluded (**Supplementary Table 28**).

or blood. This cohort contained 212 HCV-positive, 117 HBV-positive and 150 non-virus cases. The US cohort contained European-ancestry (55%), Asian (defined as US-Asian hereafter; 16%) and African-American (12%) cases. The clinical backgrounds for this cohort are shown in **Supplementary Table 1**.

The exons and surrounding noncoding genomic regions of protein-coding genes were captured in 452 pairs of tumor and non-cancerous liver tissues. Oncoviral genomes, including for HBV, human papillomavirus (HPV-16 and HPV-18) and human T-lymphotrophic virus 1 (HTLV1) (91 kb in total; **Supplementary Table 2**), were also captured in 198 cases. Whole-genome sequencing was conducted in 22 HCC pairs, including 9 exome-sequenced cases, and targeted resequencing of liver cancer genes was carried out for 38 cases. To minimize multicenter study bias due to differences in exome sequencing platform or data analysis pipeline, we optimized the somatic mutation detection algorithms and filtering conditions for three centers using Japanese cohort samples. High concordance (>87%) with a validation rate of >97% in somatic mutation detection was achieved, and substitution patterns among the three centers were consistent (**Supplementary Figs. 1** and **2**). We also confirmed that similar mutation spectra were observed in the same cases in whole-genome sequence and whole-exome sequence (**Supplementary Fig. 3**).

The average mutation rate was 2.8 mutations per megabase, and T>C and C>T substitutions were dominant in this cohort (**Supplementary Fig. 4**). Eight (1.7%) outlier tumors harboring more than 4.3 mutations per megabase showed substitution patterns distinctive from those of other cases and had somatic nonsense or missense mutations in mismatch repair (*MSH3*, *MSH4*, *MSH5* and *MSH6*), DNA polymerase (*POLA1*, *POLK*, *POLE* and *POLL*) or nucleotide excision repair (*ERCC1* and *ERCC2*) genes (**Supplementary Fig. 5**).

Panoramic view of ploidy, copy number and virus integration

We evaluated copy number alteration (CNA) by comparing the sequence depth for paired samples and allelic imbalance in the captured area (**Supplementary Fig. 6**). This digital assessment of CNA and allelic imbalance was consistent with SNP array data in cases analyzed by both methods (**Supplementary Fig. 7**). We also imputed deviation in the allele frequency of heterozygous single-nucleotide variation to predict the tumor purity and ploidy for each sample (H.U., S.Y., K.T. and H.A., unpublished data). A large fraction of cases (28.9%) represented whole-genome duplication with gross chromosomal loss (average ploidy was 3.87, and the average number of CNAs was 11.58) (**Supplementary Fig. 8**), whereas the remainder showed more stable copy number status (average ploidy was 2.08, and the average number of CNAs was 7.56). Tetraploidy was more frequently observed in higher-grade tumors ($P = 0.039$, Fisher's exact test; **Supplementary Fig. 9**).

We observed recurrent arm-level gains (1q, 5p, 6p and 8q) and losses (1p, 4q, 6q, 8p and 17p), as previously described for HCC[9] (**Supplementary Fig. 10**). Recurrent focal amplifications were detected in 25% of cases, including for *TERT* and *CCND1-FGF19*. Homozygous deletions were less frequent events (detected in 17.4% of cases). Recurrent homozygous deletion was observed for 28 genes, including *CDKN2A-CDKN2B*, *MAP2K3* and *PTEN* (**Supplementary Figs. 11** and **12**).

Using paired-end reads mapped to the HBV viral and human genomes, respectively, we detected 628 HBV virus integrations in 68 HBV-positive cases from which viral genomes were captured (9.2 integrations per case) (**Supplementary Table 3**), reflecting a detection rate that was 2–4 times more sensitive than in previous whole-genome sequencing studies[10,11]. Genes close to (less than 10 kb away from) the recurrent HBV integrations included *TERT* ($n = 17$ cases), *KMT2B* (*MLL4*; $n = 6$ cases), and *ALOX5*, *ZFPM2*, *SENP5*, *MYO19* and *RGS22* ($n = 2$ cases each). Recurrent non-genic HBV integrations were observed near the centromere, especially on chromosomes 1p, 8p and 10q. A significant fraction of HBV integrations were colocalized with (less than 500 kb away from) DNA copy number breakpoints (10.7%; $P < 1 \times 10^{-5}$, randomization test) (**Supplementary Figs. 13** and **14**). Despite intimate association between HBV genome integration and CNA breakpoints, the frequency of CNA was not different among the viral subtypes ($P = 0.29$, ANOVA test; **Supplementary Fig. 15** and **Supplementary Table 4**).

Multiple types of *TERT* genetic alteration in HCC

Somatic mutations in the transcriptional regulatory region of the *TERT* gene have been reported in a range of cancers, including HCC[12,13]. By combining captured noncoding sequence data with capillary sequencing validation, we detected *TERT* promoter mutations in 254 cases of the 469 cases analyzed (54% in total). The frequency of these mutations was highest in HCV-positive cases (121/188; 64%), with lower frequencies in non-viral cases (88/149; 59%) and HBV-positive cases (44/120; 37%) (**Supplementary Table 5**). As reported[13], the mutation located 124 bp upstream of the ATG start site (c.–124C>T, on the opposite strand; 93%) was more frequent than the c.–146C>T (4.3%) and c.–57A>C (1.6%) mutations (**Supplementary Table 6**).

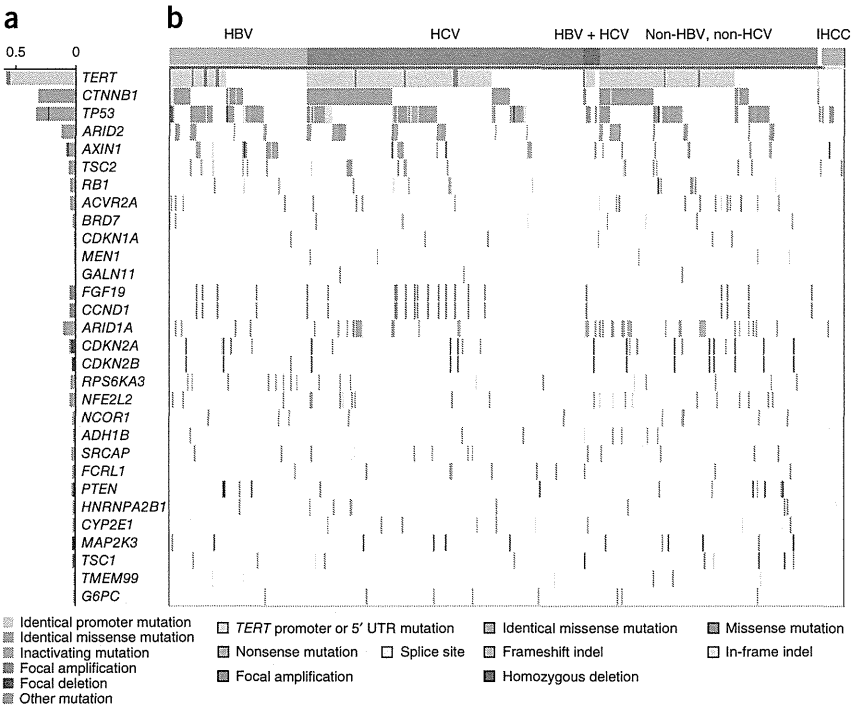

Figure 2 Significant cancer driver genes in HCC. An overview of significant driver genes in HCC. Shown are genes with statistically significant mutations or focal CNAs (**a**) and their alterations in each sample classified by the status of hepatitis virus infection (**b**). Genes were sorted by significant *q* value (**Supplementary Note**).

Additionally, *TERT* focal amplification was detected in 6.7% of the cases in total, and integration of the HBV genome in the *TERT* locus was observed in 22% of HBV-positive samples for which integration was analyzed. *TERT* promoter mutations were mutually exclusive with HBV genome integration in the *TERT* locus in integration-analyzed HBV-positive samples and were almost mutually exclusive with *TERT* focal amplifications, both of which were considered to cause higher *TERT* expression[14] (**Fig. 1**). Alterations of *ATRX* have also been reported to induce telomerase-independent telomere maintenance[15], Altogether, more than 68% of the HCC cases had alterations in either *TERT* or *ATRX*, representing the most frequent molecular event reported (**Supplementary Table 5**). In contrast, no *TERT* promoter mutations were detected in 13 IHCC cases (**Fig. 2**). *TERT* promoter mutations significantly co-occurred with WNT pathway gene alterations, such as *CTNNB1*, *AXIN1* or *APC*, in HCV-positive and non-virus cases, suggesting a cooperative oncogenic activity between *TERT* promoter mutation and the WNT pathway[16] in these subgroups (**Fig. 1**).

Significantly altered genes in HCC

To identify significantly altered genes in HCC, we used a combination of MutSigCV[17], an aggregated somatic alteration method that aggregates somatic substitutions, short indels, homozygous deletions and focal amplifications, and an inactivation bias method that calculates inactivating mutation bias (**Supplementary Fig. 16**, **Supplementary Tables 7–10** and **Supplementary Note**). Furthermore, we eliminated mutated genes that exhibited sequencing center bias and subclone bias as sources of possible false discovery (**Supplementary Tables 11** and **12**). These steps led to a final list of 30 candidate driver genes (**Fig. 2**, **Supplementary Fig. 17** and **Supplementary Tables 13–15**), including 13 that were not recurrently mutated in previous cohorts[18–20] (**Supplementary Table 16**). These 13 genes included *BRD7*, a component of the SWI/SNF nucleosome-remodeling machinery, and *MEN1*, a putative tumor suppressor somatically mutated in neuroendocrine tumors—neither of which has been reported in HCC. Mutations in *TSC2*, *SRCAP* and *NCOR1* have been reported as singletons in other

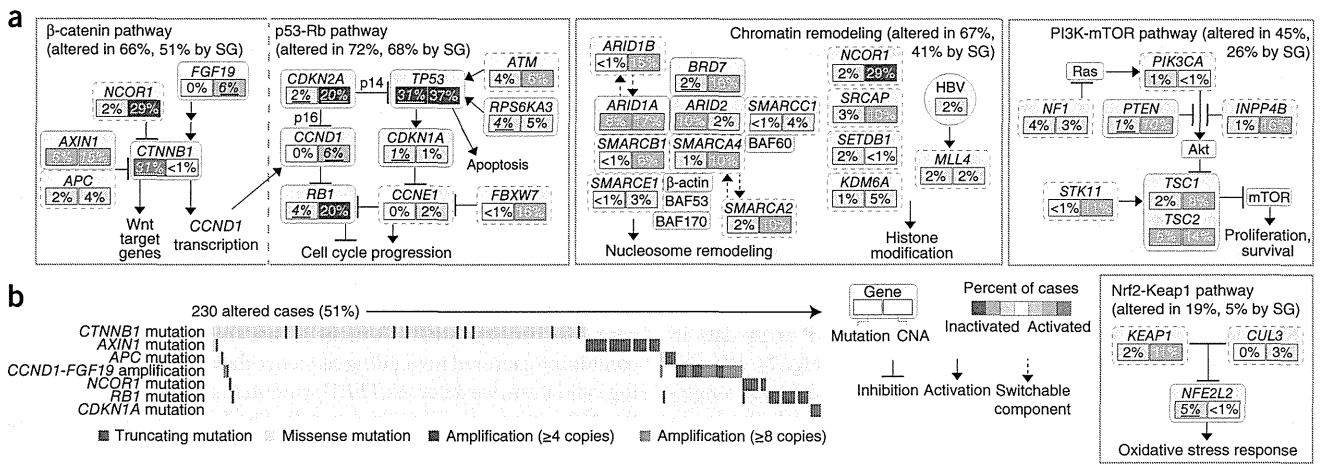

Figure 3 Oncogenic network in HCC. (**a**) Major signaling pathways involving genetic alterations in HCC. Key genes in each pathway are indicated by rectangles, with the percentages of somatic mutations and CNAs shown in the left and right portions of each rectangle, respectively. Significantly altered genes (SG; MutSigCV, $P < 0.05$ or GISTIC, q value < 0.1; percentages are underlined for alterations meeting either criterion) are bounded by solid lines, whereas other key genes in each pathway are bounded by dashed lines. (**b**) Mutual exclusivity plot of genes relevant to the WNT signaling pathway. The plot indicates that somatic mutations in WNT-related genes might contribute to the activation of WNT signaling in over half of all HCCs.

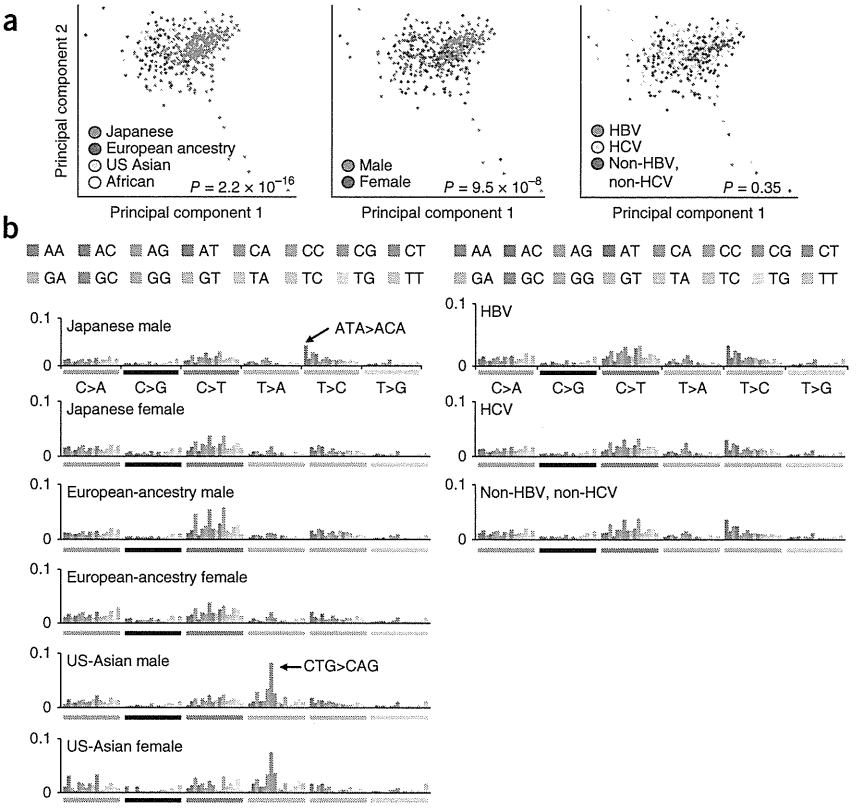

Figure 4 Somatic substitution patterns were associated with ancestry. (**a**) Principal-component analysis of the 96 substitution patterns in the HCC genome by ancestry group (left), sex (middle) and hepatitis virus group (right). (**b**) Average frequency of the 96 substitution patterns in each sample group (ancestry group, sex and virus group). The top legend shows the bases immediately 5′ and 3′ to each substitution. The *y* axis indicates the frequency of the 96 substitution patterns.

studies, but these genes were shown here to be significantly mutated. Some of the difference in results might be attributed to the greatly increased statistical power with our 503-case population, but some of the difference might also reflect contribution from the ancestry composition of the cohorts in this study. Several genes demonstrated differences in mutational frequency among virus subtypes (**Fig. 2b** and **Supplementary Table 17**). *AXIN1* was more frequently mutated in HBV-positive cases in comparison with HCV-positive and non-virus HCC (P = 0.0055, Fisher's exact test), indicating that different viral etiologies might activate WNT signaling in distinct ways. *ARID1A* was more frequently altered in non-virus cases (P = 0.009).

Alterations of drug target kinases were rarely found in HCC; low-level recurrent mutations of *FGFR2* (mutated in 1.8% of cases), *KIT* (1.3%), *FGFR3* (0.9%), *FGFR1* (0.9%), *JAK1* (0.9%) and *EGFR* (0.4%) and focal amplification of *MET* (0.5%) were detected. The specific mutations in these receptor tyrosine kinases were not generally observed in other cancers, with the exception of two *JAK1* mutations (encoding p.Ser703Ile and p.Leu910Pro substitutions), which were previously observed in a liver cancer sequencing study[20]. The liver has a central role in many metabolic processes. Our study identified recurrent mutations of metabolic enzyme genes in HCC (**Fig. 2b** and **Supplementary Tables 7** and **13**). These included *CYP2E1* (2.0%); *ADH1B* (1.8%), encoding alcohol dehydrogenase 1B; and *G6PC* (1.8%), encoding a glucose-6-phophatase catalytic subunit, whose aberrations could be linked to metabolomic changes in HCC.

Significant oncogenic pathways in HCC

Oncogenic pathways were further explored by aggregating the alterations of each gene within a particular pathway (**Fig. 3a**).

TP53-RB pathway. Inactivation of the tumor-suppressor TP53-RB pathway was a consistent theme in HCC. *TP53* mutations were observed in 31% of tumors, and two genes encoding p53-activating kinases, *ATM* and *RPS6KA3*, were also recurrently mutated. The *RB1* gene was mutated in 4.4% of cases. The *CDKN2A* gene encoding the RB regulator $p16^{INK4A}$ was subject to frequent focal homozygous deletion, and the p53 target and RB regulator *CDKN1A* (encoding $p21^{CIP1}$) was significantly mutated. Overall, 72% of cases had alterations in component genes of one or both of these pathways.

WNT pathway. In addition to activating *CTNNB1* mutations, inactivating mutations were frequently observed in WNT regulators, including *AXIN1* and *APC*. *CCND1* is a key downstream target of WNT signaling[21], and *FGF19* has been shown to activate *CTNNB1* transcriptional functions[22]. Mutual exclusivity of *CTNNB1*, *AXIN1* and *APC* mutations and *CCND1-FGF19* amplification supports the functional role of these genes in altering WNT signaling (**Fig. 3b**). Overall, 66% of HCCs showed WNT pathway–related alterations.

Chromatin and transcription modulators. A large proportion of the genes on the list of significantly mutated genes encoded chromatin modulators or transcriptional regulators. Frequent alterations in *NFE2L2*, encoding a transcriptional regulator that activates antioxidant and cytoprotective target genes[23], and its negative regulators *KEAP1* and *CUL3* (ref. 24) were noted. Also mutated were the nucleosome remodelers *ARID1A*, *ARID2* and *BRD7*, with CNAs and mutations in six additional members of the SWI/SNF complex (**Fig. 3a**), *SRCAP* and the transcriptional corepressor *NCOR1*, both of which have roles in steroid receptor–mediated transcription. These genes displayed primarily inactivating frameshift and nonsense mutations that suggest a tumor-suppressor gene function in HCC (**Supplementary Fig. 18** and **Supplementary Table 9**). *NCOR1* has been shown to directly suppress *CTNNB1* function[25] and exhibits mutual exclusivity for mutations with other WNT pathway genes (**Fig. 3b**). *SRCAP* encodes an Snf2-related CREBBP activator in several pathways, including NOTCH[26] and steroid receptors[27]. Truncating *SRCAP* mutations cause a rare hereditary disease with developmental defects and early-onset tumor formation[28,29], highlighting its potential function as a tumor-suppressor gene.

mTOR-PIK3CA pathway. Recurrent inactivating mutations in *TSC1-TSC2* and activating mutations and copy gain in *PIK3CA* were observed (**Fig. 3a**). Other modulators involved with this pathway, such as *NF1*, *PTEN*, *INPP4B* and *STK11*, were also affected, and, in total, 45% of cases had alterations in the mTOR-PIK3CA pathway. Somatic *TSC1* mutation was reported as a potential predictive biomarker of an mTOR inhibitor[30], and *TSC1*-mutated HCC cell lines showed

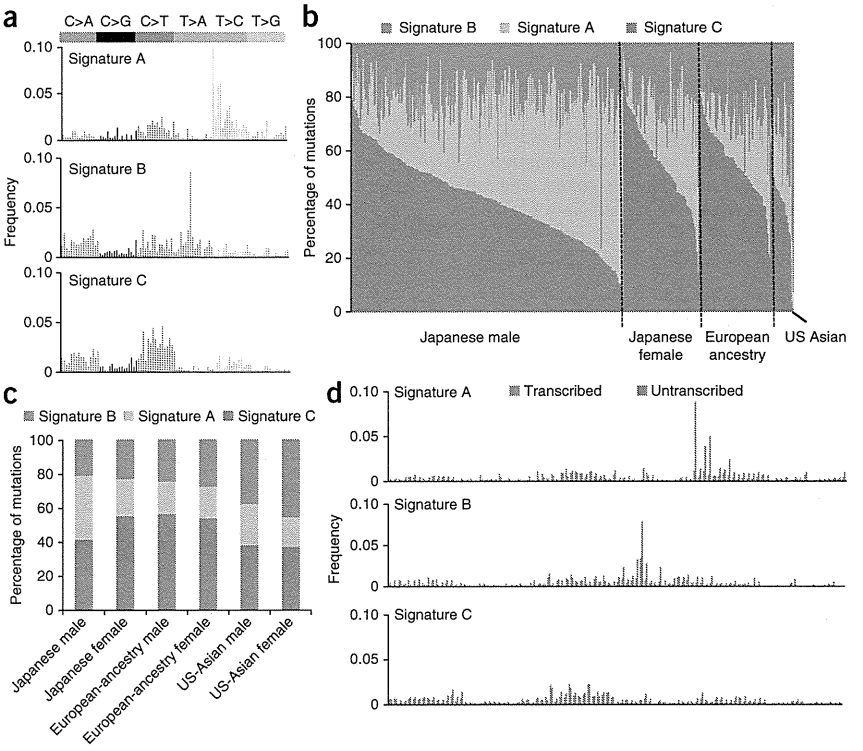

Figure 5 Ancestry-specific mutational signatures with transcriptional strand bias in the HCC genome. (**a**) The 3 mutational signatures in the HCC genome are shown according to the frequencies of 96 substitution types. The *y* axis indicates the frequency of each of the 96 substitution patterns. (**b**) Contribution of the three mutational signatures to each tumor. The *y* axis indicates the percentage of mutations comprised in each signature. The *x* axis indicates tumors classified in each ancestry group and by sex. (**c**) Contribution of the three mutational signatures to tumors from each ancestry group and sex. The *y* axis indicates the percentage of mutations comprised in each signature. (**d**) Transcriptional strand bias in mutational signatures. Each signature is displayed with 192 mutation patterns based on the 96 substitution types with transcriptional strand information. The mutation types are shown on the *x* axis, and the *y* axis indicates the frequency of each of the 192 mutation types contributing to each signature.

higher sensitivity to an mTOR kinase inhibitor (BEZ235) in comparison to cell lines with wild-type *TSC1* (**Supplementary Fig. 19**).

To identify networking among the oncogenic pathways in HCC, we developed a pathway compression algorithm and applied it to the significantly altered genes. We identified 11 core oncogenic network modules in HCC (**Supplementary Table 18**). To visualize these modules in the context of a biological network, we constructed a schematic view of the modules and the additional nodes that can connect them (**Supplementary Fig. 20**). The nodes were typically classified into two types; one type was closely connected to neighboring nodes (with higher value for centrality; **Supplementary Table 19**) and the other type had long-range edges that reached distant nodes, which can be used to measure the effect of each module alteration on the total network. Further comparison of the association between these module alterations and background clinical factors showed that the mTOR module was significantly different ($P < 0.05$, Cochran-Mantel-Haenszel test) in Asian and European-ancestry populations with respect to mutational frequencies (**Supplementary Fig. 21**).

Ancestry-dependent diversity in HCC mutation signatures

Somatic mutation patterns in human cancer are closely associated with epidemiological factors[31–34]; however, their association with ancestry remains unexplored. We integrated genomic data from an additional 105 HCC cases sequenced by TCGA along with the 503 cases sequenced by us (**Supplementary Table 1**) and compared somatic substitution patterns according to epidemiological data and ancestry group. Because mutation patterns in hypermutated cases and IHCC were distinctive (**Supplementary Figs. 4** and **22**), these two groups were excluded from further mutation pattern analysis.

Principal-component analysis of the 96 possible nucleotide triplets, dependent on the bases immediately 5′ and 3′ to each substitution, showed that the constitution of substitution patterns with these triplets was significantly different by ancestry group (Japanese, US Asian and European ancestry; $P = 2.2 \times 10^{-16}$, Wilks' test) and by sex ($P = 9.5 \times 10^{-8}$) (**Fig. 4a**). Notably, substitution patterns were not significantly associated with viral status (HBV, HCV and non-viral, $P = 0.35$; **Fig. 4a** and **Supplementary Fig. 23**). T>C substitutions, particularly in an ATA context, were specifically increased in Japanese male samples, and T>A substitutions (most frequently in a CTG context) were specifically increased in US-Asian male and female samples. The distributions of the frequencies for the 96 substitution types were similar among Japanese female samples and European-ancestry male and female samples (**Fig. 4b**).

We applied non-negative matrix factorization (NMF) analysis to the 96-substitution pattern[33] and identified 3 mutation signatures (HCC signatures A–C; **Fig. 5a** and **Supplementary Fig. 24**). Each signature was composed of context-specific substitutions: HCC signature A was characterized by dominant T>C mutations, especially in an AT(A/G/T) context, whereas HCC signature B contained dominant T>A mutations, with a sharp increase in frequency for a CTG context. HCC signature C contained dominant C>T mutations, especially in an (A/C/G)CG context. The distribution of these signatures was associated with ancestry and sex but not with the virus status (**Supplementary Table 20**). Among the different ancestry groups, HCC signatures A and B more frequently contributed to Japanese male (odds ratio (OR) = 2.2; $P = 0.0025$, Fisher's exact test) and US-Asian (OR = 2.5; $P = 0.00036$) cases, respectively, whereas HCC signature C was common across all ancestry groups and in both sexes (**Fig. 5b,c** and **Supplementary Fig. 25**). Remarkable differences in mutation prevalence between the transcribed and untranscribed strands were observed for T>C substitutions, especially in an AT(A/G/T) context ($P = 7.4 \times 10^{-152}$, χ^2 test), in HCC signature A and for T>A substitutions, especially in a CTG context ($P = 3.3 \times 10^{-8}$), in HCC signature B (**Fig. 5d**). These significant strand biases imply the involvement of transcription-coupled repair, which is tightly associated with known carcinogens in other tumor types[31–34]. There was no significant association between the signature distribution and the *ALDH2* SNP rs671, which is associated with alcohol metabolism and is a more frequent genotype in the Asian population[35] (**Supplementary Table 21**).

To collect large amounts of cancer genome data from different ancestry groups and epidemiological backgrounds, we currently need to combine data from multiple institutes that apply individual analytical platforms. An important caveat in multicenter trans-ancestry analysis has been the possibility that ancestry-specific signatures can be biased by experimental or analytical differences. To avoid this potential bias, we processed the DNA from 99 Japanese HCC cases using the sequencing and analysis pipeline at the United States–based Baylor College of Medicine. Using this data set from a single center, we replicated exactly the same signatures in each population (**Supplementary Fig. 26**). We also examined the distribution of signatures among three centers using Japanese male samples and confirmed that similar distributions were seen among the three centers (**Supplementary Fig. 27**). Furthermore, we analyzed whole-genome sequencing data for 88 Chinese HCC samples[19] and successfully identified HCC signatures B and C in this independent data set (**Supplementary Fig. 28**).

Outcome analysis from mutational signatures

We analyzed the derived NMF signatures to determine whether any signature or signature component was associated with differences in outcome in the HCC cohort. NMF signature values were merged with annotated clinical data. We performed calculations using standardized signature values to control for differences in the mutation rate between the subjects. Multivariate analysis with the Cox proportional hazards model (**Supplementary Fig. 29** and **Supplementary Tables 22–26**) indicated that histological grade, HCC signature B and the interaction with HCC signature A (but not with HCC signature C) were significant predictors of outcome.

DISCUSSION

The present trans-ancestry liver cancer genome study first identified mutational signatures that are independent of hepatitis virus infection and contribute more to the Asian cases than to ones of European ancestry (**Supplementary Tables 27**). One signature, characterized by AT>AC mutations, was predominant in Japanese males, whereas the other, featuring CTG>CAG mutations, was found more frequently in tumors from Asians living in the United States. These correlations may highlight deeper intra-ancestry diversity and/or environmental contributions, and sex bias might further affect downstream target genes and molecular features in HCC[36]. As several genetic loci are associated with individual HCC risk together with HBV and/or HCV infection[37,38], somatic and germline genome interaction might also be important to consider. Notably, these signatures were not evident in IHCC for Japanese cases (data not shown), suggesting that they are unique properties of HCC. The causes of these signatures remain unknown, but skewed transcriptional strand biases in characteristic sequence contexts strongly imply the presence of specific, previously unexplored mutational processes, which profoundly influence tumor genome constitution and behavior.

With 503 cases, this study is the largest liver cancer genome analysis thus far, enabling the formation of a more thorough picture of the mutational landscape of HCC than ever before. In addition to identifying a large number of significantly mutated genes, we have also identified recurrent alterations of 9 of the 14 core genes making up the SWI/SNF complex. We also find a combination of hotspot *TERT* promoter and *ATRX* mutations, along with focal amplification and virus genome integration in the *TERT* locus, in more than 68% of HCC cases regardless of virus subtype. These findings show that *TERT* is a central driver gene and a promising molecular target[39] in HCC. The targeting of high-prevalence mTOR-PIK3CA pathway activation and antiproliferative activity in HCC cells by chemical inhibition should also offer new therapeutic opportunities. In addition, newly identified alterations in the chromatin-remodeling complex and metabolic enzymes are expected to be associated with cancer-specific epigenetic and metabolomic features.

URLs. DNAcopy, http://www.bioconductor.org/packages/2.13/bioc/html/DNAcopy.html; R software, http://www.R-project.org/; R survival package, http://CRAN.R-project.org/package=survival/; HGSC Mercury analysis pipeline, https://www.hgsc.bcm.edu/software/mercury; GRCh38 human reference genome, http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/; BWA2, http://bio-bwa.sourceforge.net/; GATK4, http://www.broadinstitute.org/gatk/.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Sequence data have been deposited in the European Genome-phenome Archive (EGA) under accession EGAS00001000389, the ICGC database (http://www.icgc.org/) and the database of Genotypes and Phenotypes (dbGaP) under accession phs000509.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

This study was supported by Grants-in-Aid from the Ministry of Health, Labour and Welfare of Japan for the third-term Comprehensive 10-Year Strategy for Cancer Control, grants from the US National Human Genome Research Institute (NHGRI; 5U54HG003273) and National Cancer Institute (NCI; HHSN261201000053C and P30 CA125123), the Program for Promotion of Fundamental Studies in Health Sciences from the National Institute of Biomedical Innovation (NIBIO, Japan) and the National Cancer Center Research and Development Funds (23-A-8, Japan). The National Cancer Center Biobank is supported by the National Cancer Center Research and Development Fund, Japan. The supercomputing resource SHIROKANE was provided by the Human Genome Center at the University of Tokyo (http://sc.hgc.jp/shirokane.html).

AUTHOR CONTRIBUTIONS

Study design: Y.T., K.T., K.R.C., H.U., M.K., D.A.W., H.A. and T.S. Sequencing data generation: K.T., D.M.M., F.H., H. Doddapaneni, H. Dinh, Y.A., K.G., K.W., M.-C.G., T.U., S.O., N.O., M.W. and Y.Z. Data analysis: Y.T., K.T., K.R.C., H.U., M.K., S.T., L.A.D., B.L.S., E.S., S.Y., H.N., M.L., N.H., K.W., K.G., M.D., G.N., D.A.W. and T.S. Statistical analysis: Y.T., K.R.C., H.U., K.T., C.J.C., M.K., S.T. and S.Y. Molecular analysis: Y.A. and T.S. Sample acquisition and clinical data collection: M.-C.G., K.S., Y.M., J.A.G., H.O., A.H., J.S., R.C., J.G., S.I., M.T., T.O., N.K., T.K., T.T. and M.F. Manuscript writing: Y.T., K.T., K.R.C., H.U., C.J.C., L.A.D., B.L.S., M.K., D.A.W., H.A. and T.S. Project oversight: D.A.W., R.A.G., H.A. and T.S.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Jemal, A. *et al.* Global cancer statistics. *CA Cancer J. Clin.* **61**, 69–90 (2011).
2. Forner, A., Llovet, J.M. & Bruix, J. Hepatocellular carcinoma. *Lancet* **379**, 1245–1255 (2012).
3. El-Serag, H.B. Epidemiology of viral hepatitis and hepatocellular carcinoma. *Gastroenterology* **142**, 1264–1273 (2012).
4. Yu, J., Shen, J., Sun, T.T., Zhang, X. & Wong, N. Obesity, insulin resistance, NASH and hepatocellular carcinoma. *Semin. Cancer Biol.* **23**, 483–491 (2013).
5. Augustine, M.M. & Fong, Y. Epidemiology and risk factors of biliary tract and primary liver tumors. *Surg. Oncol. Clin. N. Am.* **23**, 171–188 (2014).
6. Tanaka, K., Sakai, H., Hashizume, M. & Hirohata, T. Serum testosterone:estradiol ratio and the development of hepatocellular carcinoma among male cirrhotic patients. *Cancer Res.* **60**, 5106–5110 (2000).
7. International Cancer Genome Consortium. International network of cancer genome projects. *Nature* **464**, 993–998 (2010).

8. Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061–1068 (2008).
9. Wang, K. *et al.* Genomic landscape of copy number aberrations enables the identification of oncogenic drivers in hepatocellular carcinoma. *Hepatology* **58**, 706–717 (2013).
10. Sung, W.K. *et al.* Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. *Nat. Genet.* **44**, 765–769 (2012).
11. Fujimoto, A. *et al.* Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. *Nat. Genet.* **44**, 760–764 (2012).
12. Killela, P.J. *et al. TERT* promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal. *Proc. Natl. Acad. Sci. USA* **110**, 6021–6026 (2013).
13. Nault, J.C. *et al.* High frequency of telomerase reverse-transcriptase promoter somatic mutations in hepatocellular carcinoma and preneoplastic lesions. *Nat. Commun.* **4**, 2218 (2013).
14. Li, Y. & Tergaonkar, V. Noncanonical functions of telomerase: implications in telomerase-targeted cancer therapies. *Cancer Res.* **74**, 1639–1644 (2014).
15. Heaphy, C.M. *et al.* Altered telomeres in tumors with *ATRX* and *DAXX* mutations. *Science* **333**, 425 (2011).
16. Hoffmeyer, K. *et al.* Wnt/β-catenin signaling regulates telomerase in stem cells and cancer cells. *Science* **336**, 1549–1554 (2012).
17. Lawrence, M.S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013).
18. Li, M. *et al.* Inactivating mutations of the chromatin remodeling gene *ARID2* in hepatocellular carcinoma. *Nat. Genet.* **43**, 828–829 (2011).
19. Guichard, C. *et al.* Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat. Genet.* **44**, 694–698 (2012).
20. Kan, Z. *et al.* Whole-genome sequencing identifies recurrent mutations in hepatocellular carcinoma. *Genome Res.* **23**, 1422–1433 (2013).
21. Tetsu, O. & McCormick, F. β-catenin regulates expression of cyclin D1 in colon carcinoma cells. *Nature* **398**, 422–426 (1999).
22. Pai, R. *et al.* Inhibition of fibroblast growth factor 19 reduces tumor growth by modulating β-catenin signaling. *Cancer Res.* **68**, 5086–5095 (2008).
23. Motohashi, H. & Yamamoto, M. Nrf2-Keap1 defines a physiologically important stress response mechanism. *Trends Mol. Med.* **10**, 549–557 (2004).
24. Zhang, D.D., Lo, S.C., Cross, J.V., Templeton, D.J. & Hannink, M. Keap1 is a redox-regulated substrate adaptor protein for a Cul3-dependent ubiquitin ligase complex. *Mol. Cell. Biol.* **24**, 10941–10953 (2004).
25. Song, L.N. & Gelmann, E.P. Silencing mediator for retinoid and thyroid hormone receptor and nuclear receptor corepressor attenuate transcriptional activation by the β-catenin–TCF4 complex. *J. Biol. Chem.* **283**, 25988–25999 (2008).
26. Eissenberg, J.C., Wong, M. & Chrivia, J.C. Human SRCAP and *Drosophila melanogaster* DOM are homologs that function in the Notch signaling pathway. *Mol. Cell. Biol.* **25**, 6559–6569 (2005).
27. Monroy, M.A. *et al.* SNF2-related CBP activator protein (SRCAP) functions as a coactivator of steroid receptor–mediated transcription through synergistic interactions with CARM-1 and GRIP-1. *Mol. Endocrinol.* **17**, 2519–2528 (2003).
28. Hood, R.L. *et al.* Mutations in *SRCAP*, encoding SNF2-related CREBBP activator protein, cause Floating-Harbor syndrome. *Am. J. Hum. Genet.* **90**, 308–313 (2012).
29. Nelson, R.A. *et al.* Floating-Harbor syndrome and intramedullary spinal cord ganglioglioma: case report and observations from the literature. *Am. J. Med. Genet. A.* **149A**, 2265–2269 (2009).
30. Iyer, G. *et al.* Genome sequencing identifies a basis for everolimus sensitivity. *Science* **338**, 221 (2012).
31. Pleasance, E.D. *et al.* A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* **463**, 184–190 (2010).
32. Pleasance, E.D. *et al.* A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature* **463**, 191–196 (2010).
33. Alexandrov, L.B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
34. Poon, S.L. *et al.* Genome-wide mutational signatures of aristolochic acid and its application as a screening tool. *Sci. Transl. Med.* **5**, 197ra101 (2013).
35. Goedde, H.W. *et al.* Population genetic studies on aldehyde dehydrogenase isozyme deficiency and alcohol sensitivity. *Am. J. Hum. Genet.* **35**, 769–772 (1983).
36. Keng, V.W. *et al.* Sex bias occurrence of hepatocellular carcinoma in Poly7 molecular subclass is associated with *EGFR*. *Hepatology* **57**, 120–130 (2013).
37. Zhang, H. *et al.* Genome-wide association study identifies 1p36.22 as a new susceptibility locus for hepatocellular carcinoma in chronic hepatitis B virus carriers. *Nat. Genet.* **42**, 755–758 (2010).
38. Kumar, V. *et al.* Genome-wide association study identifies a susceptibility locus for HCV-induced hepatocellular carcinoma. *Nat. Genet.* **43**, 455–458 (2011).
39. Harley, C.B. Telomerase and cancer therapeutics. *Nat. Rev. Cancer* **8**, 167–179 (2008).

ONLINE METHODS

DNA preparation, DNA capture and sequencing. The tissues and clinical information used in this study were obtained under informed consent and approval of the institutional review boards of each institute. DNA was extracted from liver cancer tissue and matched non-cancerous liver tissues or blood using a general protocol for genome sequencing. Exome capture was carried out using the SureSelect Human All Exon V3 or V4 plus kit depending on the samples (**Supplementary Table 28**). Preparation of sequencing libraries, DNA capture methods and Illumina sequencing were carried out as described in the **Supplementary Note**.

Mutation calling. *Mutation calling (National Cancer Center Research Institute).* Paired-end reads were aligned to the human reference genome (GRCh37) using the Burrows-Wheeler Aligner (BWA)[40] for both tumor and normal samples. Probable PCR duplications, for which paired-end reads aligned to the same genomic position, were removed, and pileup files were generated using SAMtools[41] and a program developed in house. Details on our filtering conditions are provided in **Supplementary Tables 29** and **30**.

Mutation calling (Research Center for Advanced Science and Technology). Next-generation sequencing reads were mapped to the human genome (hg19) using BWA and Novoalign independently. Reads with a minimal editing distance to the reference genome were taken to represent optimal alignments. Then, bam files were locally realigned with SRMA. Normal-tumor pair bam files were processed using an in-house genotyper (karkinos), with the variants further filtered to remove all variants observed fewer than four times or present at an allele frequency of less than 0.12 after adjustment for tumor sample purity. The variants also had to have a score of greater than Q20 (representing the root mean square of mapping quality). In addition, reads harboring the variant had to be observed in both forward and reverse orientation. If a variant was present in reads of only one orientation, we checked for strand bias using a *t* test comparing these reads to the reads without the variant, and variants with a *P* value of <0.03 for strand bias were rejected. Variants also had to be called in different sequence cycles and have at least one call that was outside of 3% of read ends. Variants could not be located within 5 bp of an indel call, and variants where the mean base quality of the supporting reads was lower than 10 on the Phred scale were removed. Germline variants having an allelic frequency of greater than 0.1 were collected for 50 normal liver exome samples and used as the panel of normal variants. Any variant that was observed in this panel with a population frequency of greater than 2% was filtered out. Finally, variants also observed in the paired normal sample with an allelic frequency of greater than 3% and sites registered in dbSNP Build 134 with validated status were removed.

Mutation calling (Baylor College of Medicine). Initial sequence analysis was performed using the Human Genome Sequencing Center (HGSC) Mercury analysis pipeline. First, the primary analysis software on the instrument produced bcl files that were transferred off the instrument to the HGSC analysis infrastructure by the HiSeq Real-Time Analysis module. Once each run was complete and all bcl files were transferred, Mercury ran the vendor's primary analysis software (CASAVA), which demultiplexed pooled samples and generated sequence reads and base call confidence values (qualities). In the next step, reads were mapped to the GRCh37 human reference genome using BWA (BWA2), producing a bam3 (binary alignment/map) file. The third step involved quality recalibration (using GATK4) and, where necessary, the merging of bam files for separate sequence events into a single sample-level bam file. Sorting of bam files, duplicate read marking and realignment to improve indel discovery all occurred at this step.

Processing the significantly mutated genes. The significantly mutated genes for this study were identified through three separate tests as described below (an aggregated somatic alteration method, MutSigCV[42] and an inactivation bias method), and the resulting gene lists were combined in a final table of significantly mutated genes (**Supplementary Table 13**).We also developed two tests to detect bias in the mutation list that could be a source of artifact (K.R.C., E.S., L.A.D. and D.A.W., unpublished data). One of these tests examined sequencing center bias, and the other examined bias in mutation allele fraction, which if consistently low would suggest that a gene was a passenger rather than a driver. Genes in the final combined table that failed these bias tests were removed from the final list of significantly mutated genes. Data from each process are shown in **Supplementary Tables 7–12**, and the steps are shown schematically in **Supplementary Figure 16**.

Aggregated somatic alteration method. We identified significantly altered genes by aggregating somatic substitutions, short indels, homozygous deletions and focal amplifications. We initially estimated the expected number of each alteration in each gene as follows.

First, the substitution rate was estimated by dividing the number of synonymous mutations in a sample by the number of synonymous sites in the genome. For each gene, the expected number of substitutions was calculated by multiplying the substitution rate by the number of nonsynonymous sites and splice sites in the gene. Because the substitution rate at CpG sites was much higher than that in other regions, the substitution rates and expected numbers of substitutions at CpG and non-CpG sites were estimated separately using the following equation:

$$\mathrm{EN} = \sum_{i=1}^{n}\left(\frac{M_{\mathrm{CG}_i} \times N_{\mathrm{CG}}}{S_{\mathrm{CG}} \times C_i} + \frac{M_{\mathrm{NCG}_i} \times N_{\mathrm{NCG}}}{S_{\mathrm{NCG}} \times C_i}\right)$$

where n is the number of samples, M_{CG_i} is the number of synonymous mutations at CpG sites in the ith sample, M_{NCG_i} is the number of synonymous mutations in non-CpG sites in the ith sample, S_{CG} is the number of synonymous sites at CpG sites in the genome, S_{NCG} is the number of synonymous sites at non-CpG sites in the genome, N_{CG} is the number of nonsynonymous sites and splice sites at CpG sites in a gene, N_{NCG} is the number of nonsynonymous sites and splice sites at non-CpG sites in a gene, C_i is the fraction of sequence coverage in the genome in the ith sample (usually the fraction of coding regions that have more than 20× sequence depth for whole-exome sequencing) and EN is the expected number of nonsynonymous and splice-site substitutions in a gene.

Second, the coding indel rate was estimated by dividing the number of coding indels in a sample by the number of coding sites in the genome. For each gene, the expected number was calculated by multiplying the coding indel rate by the coding length of a gene as follows:

$$\mathrm{EI} = \sum_{i=1}^{n}\frac{I_i \times L}{S \times C_i}$$

where I_i is the number of coding indels in the ith sample, S is the number of coding sites in the genome, L is the coding length of the gene and EI is the expected number of coding indels in a gene.

Third, as regions of focal amplification and homozygous deletion are much broader than gene regions, the number of focal amplifications and homozygous deletions affecting a gene in a sample is 0 or 1 and is not influenced by gene length. Therefore, the expected number of these events is the same for all genes. The expected numbers of focal amplifications and homozygous deletions were estimated separately by dividing the total length of the focal amplification or homozygous deletion region in a sample by the length of the genome as follows:

$$\mathrm{EA} = \sum_{i=1}^{n}\frac{A_i}{G \times C_i}$$

$$\mathrm{ED} = \sum_{i=1}^{n}\frac{D_i}{G \times C_i}$$

where A_i is the total length of focal amplifications in the ith sample, D_i is the total length of homozygous deletions in the ith sample, G is the length of the genome, EA is the expected number of focal amplifications in the gene and ED is the expected number of homozygous deletions in the gene.

Fourth, the expected number of protein-altering mutations was calculated by aggregating the expected numbers of nonsynonymous and splice-site substitutions in CpG and non-CpG sites, coding indels, focal amplifications and homozygous deletions as follows:

$$\mathrm{E} = \mathrm{EN} + \mathrm{EI} + \mathrm{EA} + \mathrm{ED}$$

where E is the expected number of protein-altering mutations in a gene.

doi:10.1038/ng.3126

Fifth, tests of the significance of each gene were performed by assuming a Poisson distribution of mutation number. Adjustment for multiple testing was performed using the Benjamini-Hochberg method[8].

Inactivation bias method. The number of missense mutations was compared to the number of inactivating mutations (nonsense, frameshift and splice site) using a χ^2 test.

Analysis of sequencing center bias. Because multiple centers participated in this study, we sought to control for the influence of differences in mutation calling strategy, which might promote a gene to significance merely because of a bias in the variant callers used. Many studies do not use multiple callers and therefore have no way to control for these biases. For each gene with more than five variants, we counted the number of subjects for whom the gene was called for each center. These counts were compared to the total number of subjects using the χ^2 test. The results of the analysis for center bias are presented in **Supplementary Table 11**.

Analysis of subclone bias. Oncogenic driver events in a given tumor should exhibit allele fractions that are roughly the same as the mean allele fraction for the entire sample for any given subject. We separated oncogenic (driver) events from recurrent passenger events by comparing the allele fraction of mutations in candidate genes to the matched mean allele fraction of the sample, across all samples in the cohort. First, the mean somatic allele fraction was calculated for each subject (AF*s*). Next, for each variant in each gene, the allele fraction for the variant (AF*g*) was compared to the AF*s* in the respective subject. We calculated the fraction of events where AF*g* was less than AF*s* and generated a *P* value using a one-sided pairwise Wilcoxon test where the alternative hypothesis was that AF*g* was less than AF*s* (always with respect to the relevant subject). The histogram of all allele fraction biases (sum(AF*g* < AF*s*)/*n*, where *n* is variant count) is shown in **Supplementary Figure 30**. Selected significantly mutated genes are plotted individually to show how known drivers are distributed by this test. Note that several tumor-suppressor genes exhibited enrichment above the average allele fractions (for example, *RB1* and *TP53*). In these cases, the genes were typically both mutated and underwent loss of heterozygosity (LOH) for the wild-type allele. The results of subclone bias testing for all genes with more than five mutations are presented in **Supplementary Table 12**.

Copy number analysis, tumor purity and adjustment of mutated allele frequency. Initial copy number estimates were obtained by comparing read depth information for tumor and normal samples using VarScan2 (ref. 43). Depth estimates were then segmented using circular binary segmentation (CBS) as implemented in the DNAcopy package in R[44]. We used the JISTIC[45] program to generate a combined copy number matrix file. The VCRome2.1 probe locations were used as marker positions for copy number analysis. We then used JISTIC to calculate the significance for copy number gains and losses. Focal amplification at the *TERT* locus was determined using the average read depth of each captured target region.

Evaluation of tumor ploidy and purity. Using bam files from normal and tumor samples, read depth was calculated for each captured target region. After normalization by the number of total reads and GC content using regression analysis, the tumor/normal depth ratio was calculated, and values were smoothed using the moving average. Copy number peaks were then estimated using wavelet analysis, and each peak was approximated using Gaussian models. Hidden Markov models (HMMs) with the calculated Gaussian peaks were constructed, and copy number peaks were linked to genomic regions. The allelic imbalance for each copy number peak was calculated on the basis of heterozygous SNPs within the assigned region, and imbalance information and peak distances were further analyzed by model fitting where the optimal solution for a copy number peak was determined using vector matching, yielding estimated copy number and tumor purity and ploidy data simultaneously. Detailed algorithms will be described elsewhere (H.U., S.Y., K.T. and H.A., unpublished data).

HBV integration analysis. *HBV integration detection.* Viral genomes (HBV, NC_003977.1; HPV-16, NC_001526; HPV-18, NC_001357; HTLV-1, NC_001436) were downloaded from NCBI and included in the reference files when reads were mapped by BWA. No read was mapped to a virus other than HBV. To achieve more precise HBV mapping, we mapped all reads to the HBV reference sequence using the *q*-gram and Smith-Waterman method. An 11-mer *q*-gram was first applied to both strands of the HBV reference, and reads with 15 or more hits were subjected to Smith-Waterman alignment. The other end of each read was mapped to the hg19 human sequence using BWA. Finally, HBV integration sites were clustered by genomic position with a window size of 300 bp (approximately equal to the library fragment size), and sites with more than three supporting reads were used in the analysis.

Randomization test of HBV integration and copy number breakpoints. The 7,891 copy number breakpoints and 1,039 HBV integration sites were detected in 70 HBV-positive samples. Coexistence of the copy number breakpoints and HBV integration sites was examined using a 500-kb window size. To show statistical significance, we performed a randomization test by switching the position of the HBV integration sites to the same number of integration sites observed in the normal sample of other cases. We repeated this switching 100,000 times to yield distributions and estimated the *P* value.

Verification of single-nucleotide variation. We validated our mutation calls for frequently mutated genes (**Supplementary Table 31**) by resequencing samples using the Ion Proton sequencer (Life Technologies). Details are provided in the **Supplementary Note**.

Sanger sequencing of the *TERT* promoter. Bidirectional sequencing of the *TERT* promoter region was completed for 519 HCC samples. PCR runs were set up using 20 ng of genomic DNA, 10 μM manually designed primers (**Supplementary Table 32**) and KAPA HiFi DNA polymerase (Kapa Biosystems, KK2612). Touchdown PCR was performed with the following parameters: an initial denaturation at 98 °C for 5 min followed by 10 cycles of 98 °C for 30 s, 72 °C for 30 s and 72 °C for 1 min (decreasing the annealing temperature by 1 °C per cycle). The reaction then continued with 30 cycles of 98 °C for 30 s, 63 °C for 30 s and 72 °C for 1 min followed by a final extension at 72 °C for 5 min. The PCR products were purified with a 1:15 dilution of Exo-SAP, diluted by 0.6× and cycle sequenced for 25 cycles using a 1:64 dilution of BigDye Terminator v3.1 reaction mix (Applied Biosystems, 4337456). Finally, reactions were precipitated with ethanol, resuspended in 0.1 mM EDTA and analyzed on ABI 3730xl sequencing instruments using the Rapid36 run module and 3xx base-caller. SNPs were identified using SNP Detector software and were validated visually with Consed.

Analysis of mutation patterns and signatures. Mutation patterns for cases with hypermutation and IHCC cases were distinct from those for HCC cases (**Supplementary Figs. 4** and **21**), and cases with a small number of mutations cannot accurately represent the frequency of mutational patterns; therefore, cases with hypermutation, IHCC cases and cases with fewer than 40 mutations were excluded from further mutation pattern analysis.

The number of each of 96 possible somatic substitution types, C>A/G>T, C>G/G>C, C>T/G>A, T>A/A>T, T>C/A>G and T>G/A>C with the bases immediately 5′ and 3′ to each substitution in coding regions, was counted for each sample. The frequency of each of these substitutions was determined by dividing each count by the total number of substitutions, and the resulting frequencies were used for principal-component analysis. Principal-component analysis was implemented using the R command prcomp with the scaling option on. We used Wilks' λ test to evaluate the significance of the mean vector differences in different populations. We applied NMF to the 96-substitution pattern using published software[13], running 1,000 iterations of NMF with each NMF run iterated until convergence was achieved (10,000 iterations without change) or until the maximum number of 1,000,000 iterations was reached. We used another published software package[14] for model selection in NMF (selecting the input number of mutational signatures). Details on model selection for our NMF analysis are provided in the **Supplementary Note** and **Supplementary Figures 31–35**.

Pathway analysis. We used gene sets from MsigDB C2.all as pathway data sets. To assign *P* values representing the enrichment of mutations in pathways, we first checked whether a gene had at least one non-silent mutation or overlapped with focal CNAs for each sample in a given pathway (gene set). If so, we referred to such a gene as a 'mutated gene' for a sample. We then computed a population frequency for pathways with at least one mutated gene in the given

pathway and divided the frequency by the total length of the unioned exons of all genes in the pathway to correct for the greater number of mutations in longer genes. This quotient was used as a test statistic. We used a bootstrapping approach to calculate *P* values. In the bootstrapping approach, we randomly selected as many genes as in the given pathway from all genes in the genome and then calculated the statistic. We repeated this sampling 2,000 times, calculating a fraction corresponding to the number of sampling results in which a statistic value was greater than or equal to the value in the observed data. This fraction was used as a *P* value.

To find intensively mutated gene modules in liver cancer tissue using the identified significantly mutated gene sets from MsigDB analysis, we used Pathway Commons[15] data for the whole unbiased human gene network and integrated the gene sets into this network. All pairs of gene relationships were weighted by how many mutated genes were shared by the two genes (shared ratio). These gene relationships constituted the gene network. The whole network was split into one large connected network and some isolated small networks. To extract gene modules, we recursively eliminated edges with low shared ratio values and distinguished into the smaller modules. Although the recursive edge elimination procedure gradually clarifies tightly connected gene modules, gene modules were rarely isolated from the whole network. Using this compression process and some additional manual curation, we finally selected ten representative modules that were intensively mutated in liver cancer tissues. We took essentially the same approach as described above to calculate *P* values for mutation enrichment and mutual exclusivity for a gene pair or combination of modules. For mutation enrichment, we used all genes in a pair of modules. For mutual exclusivity, if a module had at least one mutated gene, we referred to such a module as an 'impaired module' and computed a frequency of impaired modules for each sample.

Outcome analysis from non-negative matrix factorization signatures. NMF signature values were merged with annotated clinical data for our cohort. We performed calculations using standardized signature values to control for differences in mutational rate among the subjects. For the standardized data, the contributions of each signature within a subject summed to 1. We performed Cox proportional hazards analysis[46] using the R[44] survival package, factoring in all three signature components (signature A, signature B and signature C), age at diagnosis and histological tumor grade.

40. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
41. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
42. Lawrence, M.S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013).
43. Koboldt, D.C. *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568–576 (2012).
44. R Development Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, 2010).
45. Sanchez-Garcia, F., Akavia, U.D., Mozes, E. & Pe'er, D. JISTIC: identification of significant targets in cancer. *BMC Bioinformatics* **11**, 189 (2010).
46. Cox, D.R. & Oakes, D. *Analysis of Survival Data* (Chapman & Hall/CRC, Boca Raton, FL, 1984).

doi:10.1038/ng.3126

Higd1a is a positive regulator of cytochrome *c* oxidase

Takaharu Hayashi[a,b], Yoshihiro Asano[a,b,1], Yasunori Shintani[a], Hiroshi Aoyama[c], Hidetaka Kioka[b], Osamu Tsukamoto[a], Masahide Hikita[d], Kyoko Shinzawa-Itoh[d], Kazuaki Takafuji[e], Shuichiro Higo[a,b], Hisakazu Kato[a], Satoru Yamazaki[f], Ken Matsuoka[b], Atsushi Nakano[g], Hiroshi Asanuma[h], Masanori Asakura[g], Tetsuo Minamino[b], Yu-ichi Goto[i], Takashi Ogura[d], Masafumi Kitakaze[g], Issei Komuro[j], Yasushi Sakata[b], Tomitake Tsukihara[d,k], Shinya Yoshikawa[d], and Seiji Takashima[a,k,1]

Departments of [a]Medical Biochemistry and [b]Cardiovascular Medicine, [e]Center for Research Education, and [c]Graduate School of Pharmaceutical Science, Osaka University Graduate School of Medicine, Suita, Osaka 565-0871, Japan; [d]Department of Life Science, University of Hyogo, 3-2-1 Kouto, Kamigohri, Akoh, Hyogo 678-1297, Japan; [k]Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Agency, Kawaguchi, Saitama 332-0012, Japan; Departments of [f]Cell Biology and [g]Clinical Research and Development, National Cerebral and Cardiovascular Center Research Institute, Suita, Osaka 565-8565, Japan; [h]Department of Cardiovascular Science and Technology, Kyoto Prefectural University School of Medicine, Kamigyo-ku, Kyoto 602-8566, Japan; [i]Department of Child Neurology, National Center Hospital of Neurology and Psychiatry, National Center of Neurology and Psychiatry, Kodaira, Tokyo 187-8502, Japan; and [j]Department of Cardiovascular Medicine, Graduate School of Medicine, University of Tokyo, Tokyo 113-8656, Japan

Edited by Gottfried Schatz, University of Basel, Reinach, Switzerland, and approved December 16, 2014 (received for review October 15, 2014)

Cytochrome *c* oxidase (CcO) is the only enzyme that uses oxygen to produce a proton gradient for ATP production during mitochondrial oxidative phosphorylation. Although CcO activity increases in response to hypoxia, the underlying regulatory mechanism remains elusive. By screening for hypoxia-inducible genes in cardiomyocytes, we identified *hypoxia inducible domain family, member 1A* (*Higd1a*) as a positive regulator of CcO. Recombinant Higd1a directly integrated into highly purified CcO and increased its activity. Resonance Raman analysis revealed that Higd1a caused structural changes around heme *a*, the active center that drives the proton pump. Using a mitochondria-targeted ATP biosensor, we showed that knockdown of endogenous Higd1a reduced oxygen consumption and subsequent mitochondrial ATP synthesis, leading to increased cell death in response to hypoxia; all of these phenotypes were rescued by exogenous Higd1a. These results suggest that Higd1a is a previously unidentified regulatory component of CcO, and represents a therapeutic target for diseases associated with reduced CcO activity.

cytochrome *c* oxidase | oxidative phosphorylation | resonance Raman spectroscopy | ATP | oxygen

Cytochrome *c* oxidase (CcO) (ferrocytochrome c: oxygen oxidoreductase, EC 1. 9. 3. 1) is the terminal component of the mitochondrial electron transfer system. CcO couples the oxygen-reducing reaction with the process of proton pumping. Aerobic organisms use this reaction to form a proton gradient across the mitochondrial inner membrane, which is ultimately used by the F_oF_1-ATP synthase to produce ATP.

Mammalian CcO is composed of 13 different subunits (1) containing four redox-active metal centers, two copper sites, and two heme *a* groups. These active centers accept electrons from cytochrome *c* and sequentially donate them to dioxygen. Our group and others have extensively analyzed the link between the oxygen reduction process and proton pumping at the active centers using crystallography, resonance Raman spectroscopy, and Fourier transform infrared spectroscopy (2–4). The metal ions in the copper sites and heme groups in the active centers are individually coordinated by the surrounding amino acids. We have shown that changes in the redox state cause 3D structural changes around the active centers, which in turn leads to alteration of the proton pump mediated by specific amino acid chains that coordinate each metal group (5). Thus, binding of an allosteric regulator close to the active centers might change the efficiency of both electron transfer to oxygen and proton pumping.

Several proteins involved in oxygen supply or metabolism are transcriptionally regulated by intracellular oxygen concentration: vascular endothelial growth factor (VEGF) (6), erythropoietin (EPO) (7), and G0/G1 switch gene 2 (G0s2) for F_oF_1-ATP synthase, as we recently revealed (8). Because CcO is the only enzyme in the body that can use oxygen for energy transduction, it has been suggested that the regulatory mechanism of CcO is dependent on oxygen concentration (9–12); however, this has yet to be demonstrated. In this study, we aimed to identify a regulator of CcO driven by low oxygen concentration.

In this study, by screening for hypoxia-inducible genes, we discovered that *hypoxia inducible domain family, member 1A (Higd1a)* is a positive regulator of CcO. Furthermore, using our recently established ATP-sensitive fluorescence resonance energy transfer (FRET) probe, we demonstrated that Higd1a increased mitochondrial ATP production. We also showed that Higd1a directly bound CcO and changed the structure of its active center.

Results

***Higd1a* Expression Is Induced Early in the Response to Hypoxia.** During the first few hours of hypoxia, CcO and oxidative phosphorylation (OXPHOS) activity is activated, presumably to fully use any remaining oxygen (12). At later time points, metabolism shifts toward glycolysis. Therefore, we hypothesized that a positive regulator of CcO must be up-regulated during an early stage of hypoxia, but down-regulated when glycolysis-related genes become elevated. To identify early hypoxia responsive genes that

Significance

We identified hypoxia-inducible domain family, member 1A (Higd1a) as a positive regulator of cytochrome c oxidase (CcO). CcO, the terminal component of the mitochondrial electron transfer system, reductively converts molecular oxygen to water coupled to pump protons across the inner mitochondrial membrane. Higd1a is transiently induced under hypoxic conditions and increases CcO activity by directly interacting with CcO in the vicinity of its active center. Induction of Higd1a leads to increased oxygen consumption and subsequent mitochondrial ATP synthesis, thereby improving cell viability under hypoxia.

Author contributions: Y.A., Y. Shintani, H. Kioka, T.T., S. Yoshikawa, and S.T. designed research; T.H., Y. Shintani, H. Kioka, O.T., M.H., K.S.-I., K.T., and H. Kato performed research; H. Aoyama, M.H., Y.-i.G., T.O., M.K., I.K., Y. Sakata, T.T., and S. Yoshikawa contributed new reagents/analytic tools; T.H., Y. Shintani, S.H., S. Yamazaki, K.M., A.N., H. Asanuma, M.A., and T.M. analyzed data; and T.H., Y. Shintani, H. Kato, T.O., T.T., S. Yoshikawa, and S.T. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: takasima@cardiology.med.osaka-u.ac.jp or asano@cardiology.med.osaka-u.ac.jp.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1419767112/-/DCSupplemental.

might regulate CcO activity, we analyzed gene-expression profiles of neonatal rat cardiomyocytes, one of the most mitochondria-rich cell types, exposed to hypoxic conditions (1% oxygen for 0, 4, or 12 h). Focusing on the genes whose expression was induced more than two-fold at 4 h relative to the prestimulation stage, but then decreased by 12 h, we identified three genes (Fig. 1*A* and Fig. S1 *A* and *B*). Next, we prioritized genes that were (*i*) well conserved among eukaryotes and (*ii*) listed in MitoCarta (13); only one gene, *Higd1a*, satisfied both criteria. To analyze the endogenous expression levels of Higd1a in rat cardiomyocytes, we raised a specific antibody against Higd1a and confirmed its specificity (Fig. S2 *A* and *B*). In cardiomyocytes exposed to hypoxia, Higd1a protein levels increased gradually from 0 to 12 h and then decreased by 24 h (Fig. 1*B*). Immunofluorescence revealed that both endogenous and exogenous Higd1a localized in the mitochondria (Fig. S2*C*).

Higd1a Directly Integrates into the CcO Macromolecular Complex. Because Rcf1a, the yeast homolog of Higd1a, associates with CcO (9–11), we first tested whether mammalian Higd1a binds to CcO in vivo. Indeed, endogenous binding between Higd1a and CcO in rat cardiomyocytes was confirmed by immunocapture with an anti-Higd1a antibody (Fig. S3*A*) and verified by reciprocal coimmunoprecipitation with an anti-Cox4 antibody (Fig. S3*B*). This in vivo interaction was further validated by blue native PAGE (BN-PAGE) of mitochondrial fractions from rat cardiomyocytes (Fig. S3*C*).

Because preparation of the CcO macromolecular complex, which consists of 13 subunits, is technically demanding, it has remained unclear whether Rcf1a/Higd1a binding to CcO is direct. To address this issue, we performed an in vitro pull-down assay using highly purified bovine CcO (hpCcO), which we prepared by dissolving microcrystals used for X-ray structural analysis (14). Notably, recombinant maltose binding protein-fused bovine Higd1a (MBP-Higd1a) (Fig. S4) directly associated with hpCcO (Fig. 1*C*). Furthermore, to assess macromolecular complex formation, we performed BN-PAGE followed by immunoblotting with an antibody against Higd1a, demonstrating that recombinant Higd1a indeed integrated into hpCcO (Fig. 1*D*). With these results, we conclude that Higd1a directly associates and integrates into the CcO macromolecular complex.

Higd1a Causes Structural Changes in CcO and Influences the Active Center of Heme *a*. To explore the relevance of the interaction between Higd1a and CcO, we investigated whether recombinant Higd1a affects hpCcO enzymatic activity. Strikingly, direct addition of MBP-Higd1a to hpCcO significantly increased CcO activity to twice that of hpCcO alone or hpCcO mixed with MBP (Fig. 2*A*). This significant increase in hpCcO activity stimulated by Higd1a led us to speculate that Higd1a causes a structural change at the active centers of CcO.

Therefore, we next investigated whether Higd1a changes the intensity of the visible part of the absorption spectrum of oxidized CcO. MBP alone, used as a negative control, did not cause a significant change in the absorption spectra (Fig. S5). By contrast, MBP-Higd1a caused significant spectral changes at 413 nm and 432 nm (Fig. 2*B*), wavelengths that reflect conformational changes around the hemes in oxidized CcO (15).

To obtain further structural insights, we performed resonance Raman spectroscopy, a powerful and sensitive method for detecting kinetic structural changes that cannot be assessed by X-ray crystal structural analysis. Fig. 2*C* depicts the resonance Raman spectra of CcO with and without MBP-Higd1a, focusing on the heme structure by using 413 nm excitation. The resonance Raman band at 1,372 cm^{-1} in (a: hpCcO) and (b: hpCcO + Higd1a) is assignable to the ν_4 mode of heme and is indicative of ferric heme. After the addition of recombinant Higd1a, the resonance Raman spectra demonstrated two sets of different peaks (or band shifts) at 1,562/1,592 cm^{-1} (the ν_2 mode; a marker for the spin state of heme) (16) and 1,673/1,644 cm^{-1} (the $\nu_{CH=O}$ mode of the formyl group of heme *a*) (17). Importantly, the frequency shift of the band at 1,592 cm^{-1} to 1,562 cm^{-1} is attributable to partial conversion of heme from a low-spin to a high-spin state. In oxidized CcO, only heme *a* includes low-spin iron; therefore, heme *a*, but not heme a_3, is responsible for the band shift (16). These data suggest that the binding of Hig1a to CcO caused structural changes at heme *a*, the active center of CcO.

Higd1a Positively Regulates CcO Activity and Subsequent Mitochondrial OXPHOS. Next, we investigated whether Higd1a truly regulates CcO activity in vivo. To this end, we assessed biochemical CcO activity in rat cardiomyocytes with modified expression of Higd1a. Notably, we observed a significant decrease in CcO activity in Higd1a knockdown cells. This effect was rescued by overexpression of Higd1a, eliminating the possibility of off-target effects in the RNAi experiment (Fig. 3*A*, *Left*). Moreover, overexpression of Higd1a alone increased the basal CcO activity (Fig. 3*A*, *Right*). These data suggest that Higd1a is an endogenous and positive regulator of CcO.

To assess the effect of Higd1a on cellular respiration, we continuously measured the oxygen consumption rate (OCR) using a XF96 Extracellular Flux Analyzer (Seahorse Bioscience).

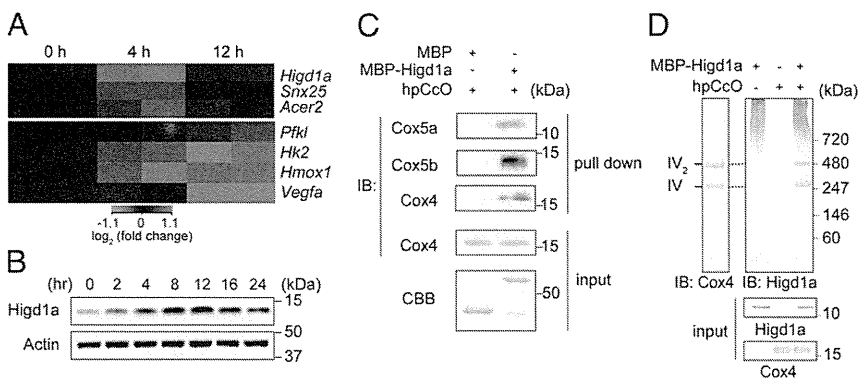

Fig. 1. Hypoxia-inducible Higd1a directly binds to highly purified cytochrome *c* oxidase (hpCcO). (*A*) Heat map of three genes (*Upper*) identified as relatively rapid and transiently induced in response to hypoxia in rat neonatal cardiomyocytes, compared with genes known to be hypoxia inducible (*Pfkl*, *Hk2*, *Hmox1*, and *Vegfa*) (*Lower*). (*B*) Expression of the Higd1a protein was elevated in response to hypoxia. (*C*) In vitro pull-down assay with amylose resin revealed direct binding between MBP-Higd1a and the hpCcO from bovine heart. Loading controls for the hpCcO and MBP-fusion proteins are shown in immunoblots for anti-CcO subunits and CBB staining, respectively. (*D*) MBP-Higd1a directly integrates into hpCcO. Mixed MBP-fusion proteins and hpCcO containing 0.2% *n*-decyl-β-D-maltoside (DM) were resolved by blue native PAGE (BN-PAGE), followed by immunoblotting with anti-Cox4 to detect CcO and anti-Higd1a to detect Higd1a.

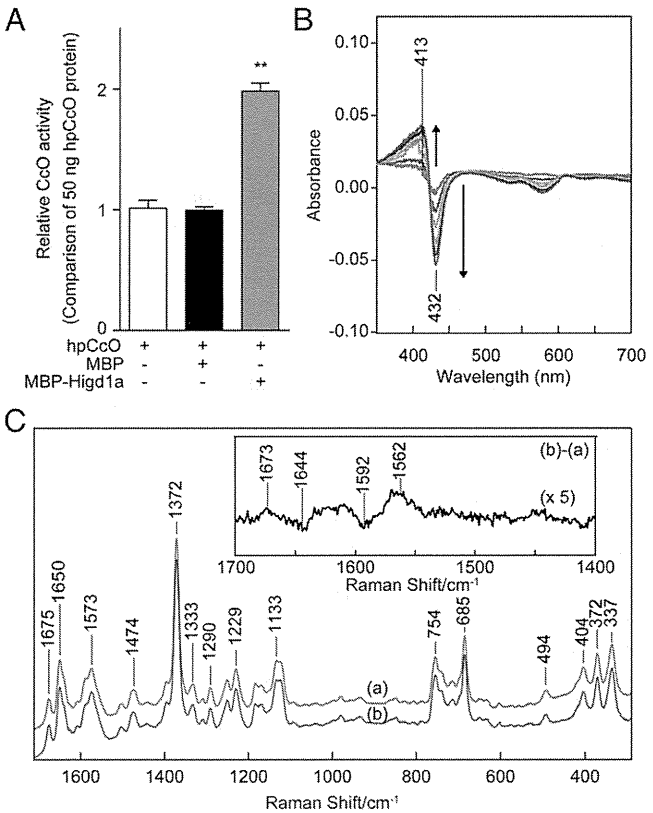

Fig. 2. Higd1a regulates CcO activity through the structural change of the active center in CcO. (*A*) CcO activity of hpCcO and hpCcO with either recombinant MBP or recombinant MBP-Higd1a. MBP-Higd1a causes an increase in CcO activity by almost twofold. Data represent the means ± SEM of five individual experiments. $^{**}P < 0.01$, compared with MBP. (*B*) The difference in absorption spectra between MBP-Higd1a and oxidized hpCcO. MBP-Higd1a caused spectral changes at 413 and 432 nm. Intensity changes of oxidized hpCcO spectra are plotted at 1 min (red), 5 min (brown), 10 min (dark yellow), 15 min (green), 20 min (light blue), 25 min (blue), and 30 min (purple) after adding MPB-Higd1a. (*C*) Resonance Raman spectra of oxidized hpCcO at 0–5 min [spectrum (a)] and oxidized hpCcO mixed with MBP-Higd1a at 0–5 min [spectrum (b)]. The *Inset* shows the difference of the spectra [(b) − (a)].

Knockdown of Higd1a caused a significant decrease in both basal (Fig. S6*A*, *Left*) and maximum OCR, and these effects were rescued by exogenous expression of Higd1a (Fig. 3*B*, *Left*). Moreover, overexpression of Higd1a significantly increased both basal and maximum OCR (Fig. S6*A*, *Right* and Fig. 3*B*, *Right*).

Because the electron transport chain creates a proton gradient that drives F_oF_1-ATP synthase (complex V), ATP production is the overall outcome of mitochondrial OXPHOS. To determine whether modulation of CcO activity by Higd1a affects ATP production, we performed the mitochondrial activity of streptolysin O permeabilized cells (MASC) assay, a sensitive means of measuring the mitochondrial ATP production rate in semi-intact cells (18). Indeed, Higd1a knockdown caused a significant decrease in the ATP production rate relative to the control (Fig. 3*C*), whereas overexpression of Higd1a increased it (Fig. 3*D*). These results suggest that Higd1a modulates mitochondrial OXPHOS through CcO.

Higd1a Protects Cardiomyocytes Under Hypoxic Conditions by Increasing ATP Production. We reasoned that endogenous induction of Higd1a by hypoxia serves to maintain ATP production in mitochondria to the greatest extent possible when oxygen supply is limited. The intramitochondrial matrix ATP concentration ($[ATP]_{mito}$) reflects mitochondrial ATP production far more sensitively than the cytosolic ATP concentration (8). Therefore, we next assessed the effect of Higd1a on ATP production in living cells using the FRET-based mitochondrial ATP biosensor Mit-ATeam (19). First, we examined the effect of KCN, an inhibitor of CcO. KCN significantly reduced the $[ATP]_{mito}$ (Fig. S7), suggesting that Mit-ATeam provides an effective means to monitor the functional consequences of changes in CcO activity. We then confirmed that hypoxia caused a gradual decline in $[ATP]_{mito}$. Overexpression of Higd1a alleviated the decline in $[ATP]_{mito}$ during hypoxia, whereas knockdown of Higd1a accelerated the decrease in $[ATP]_{mito}$ relative to the control (Fig. 4*A*).

The yeast homolog Rcf1 plays a role in respiratory supercomplex stability, and the same is true for Higd2a, but not Higd1a (11). We investigated whether Higd1a affects respiratory supercomplex stability Higd1a-knockdown or -overexpressing cells. As shown in Fig. S8, there was no significant change in the abundance or composition of the respiratory supercomplex, suggesting that the effect of Higd1a described above is not a result of changes in supercomplex stability.

Finally, to test whether the effects of Higd1a on mitochondrial ATP synthesis affected overall cell viability, we analyzed the viability of cardiomyocytes subjected to hypoxia. Under hypoxic conditions, Higd1a-knockdown cells showed a significant increase in cell death, and this effect was rescued by exogenous expression of Higd1a (Fig. 4*B* and Fig. S9*A*). In addition, overexpression of Higd1a alone

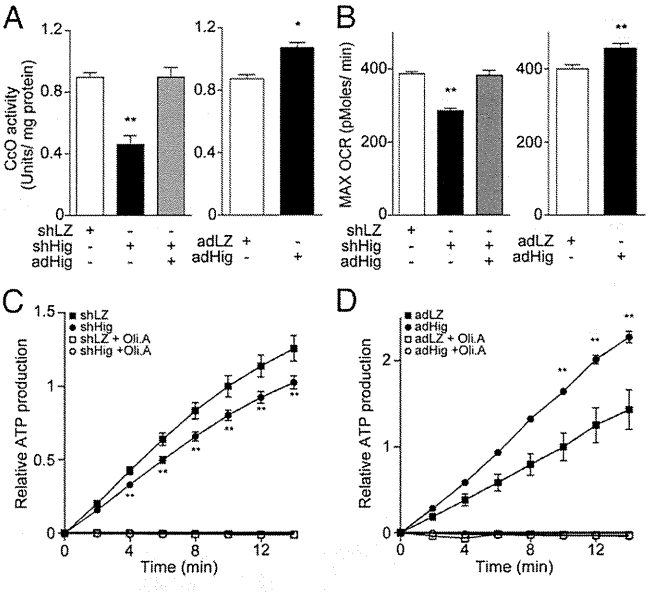

Fig. 3. Higd1a positively modulates mitochondrial respiration by altering CcO activity. (*A*, *Left*) Mitochondrial fraction from rat cardiomyocytes expressing shLacZ (shLZ), shHigd1a (shHig), or both shHig and adHigd1a (adHig) were subjected to the CcO activity assay. (*Right*) CcO activity was measured in cardiomyocytes treated with either adLacZ (adLZ) or adHig. Data represent the means of four individual experiments. (*B*, *Left*) The maximum oxygen consumption rate (max OCR) in rat cardiomyocytes transfected with the indicated adenovirus was measured after treatment with oligomycin A and fluorocarbonyl cyanide phenylhydrazone (FCCP). Knockdown of Higd1a resulted in a significant decrease in max OCR, which was rescued by exogenously expressed Higd1a. (*Right*) Overexpression of Higd1a significantly increased max OCR compared with the cells with adLZ ($n = 20$ for each group). (*C*) The relative ATP production rate of cardiomyocytes treated with shLZ or shHig was measured by the MASC assay ($n = 6$). A numerical value of ATP production at 10 min in shLZ groups is regarded as 1.0. (*D*) The relative ATP production rate of cardiomyocytes treated with adLZ or adHig was measured by MASC assay ($n = 5$). A numerical value of ATP production at 10 min in adLZ groups is regarded as 1.0. Data represent the means ± SEM; $^{*}P < 0.05$, $^{**}P < 0.01$.

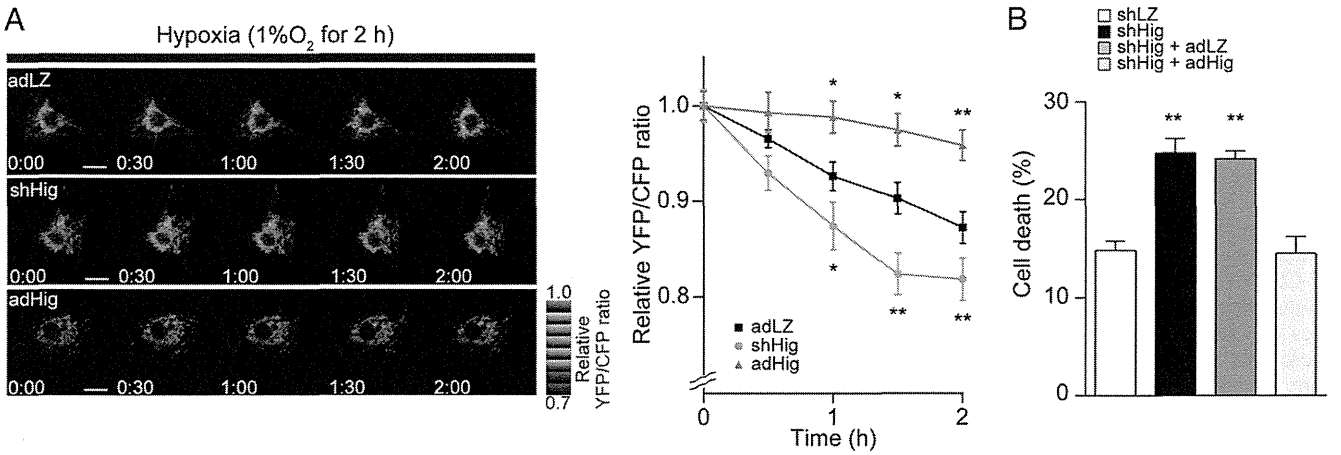

Fig. 4. (*A*) Representative sequential YFP/CFP ratiometric images of Mit-ATeam fluorescence in cardiomyocytes expressing corresponding adenovirus during hypoxia ($n = 12$ for adLZ, $n = 23$ for shHig, $n = 18$ for adHig). All of the measurements were normalized to the ratio at time 0 and compared between adLZ and adHig or shHig. (Scale bar, 20 μm.) (*B*) Cell death of cardiomyocytes treated with shHig was significantly increased compared with the control, which was rescued by addition of adHig under hypoxic conditions for 24 h ($n = 12$ for each group). Data represent the means of three independent cultures, ± SEM; $*P < 0.05$, $**P < 0.01$, compared with control (adLZ or shLZ).

increased cellular tolerance to hypoxia (Fig. S9*B*). On the basis of these findings, we conclude that Higd1a positively regulates CcO activity and subsequently increases mitochondrial ATP production, thereby protecting cardiomyocytes against hypoxia.

Discussion

In this study, we demonstrated that recombinant Higd1a produced in *Escherichia coli* was incorporated into CcO complex purified from bovine heart. The data suggest that Higd1a directly bound to the already assembled CcO complex and increased its activity. Together with the fact that Higd1a expression was rapidly increased by hypoxia, this observation indicated that Higd1a is a positive regulator of CcO that preserves the proton-motive force under hypoxic cellular stress. Physiologically, Higd1a preserved ATP production in healthy cardiomyocytes under hypoxic conditions, which protected them from an energy crisis leading to cell death.

We demonstrated that Higd1a incorporated into the CcO complex and increased its activity. It remains unclear which part of CcO is essential for this change. Higd1a binding may affect the interaction of cytochrome *c* with CcO, modulate internal electron/proton transfer, or modify K_d/K_m for O_2 binding to Cu_B/heme a_3. In fact, the resonance Raman spectroscopy experiment provided us with a clue to this question. First, we discovered that Higd1a markedly shifted the maximum Soret peak around 413 nm absorption, suggesting the occurrence of structural changes in heme that are usually observed during the reduction and oxidation process of CcO. This shift in absorbance prompted us to perform resonance Raman analysis at 413 nm excitation, a powerful tool for investigating the structure of heme and its vicinity. Higd1a induced a frequency shift of the band at 1,592 cm^{-1} to 1,562 cm^{-1} and 1,673/1,644 cm^{-1}; the former frequency is attributed to partial conversion of heme from a low-spin to a high-spin state. In oxidized CcO, only heme *a* includes low-spin iron (16); therefore, heme *a*, but not heme a_3, is responsible for this band shift.

X-ray structural and mutational analyses for bovine heart CcO have demonstrated that protons are pumped through the hydrogen-bond network across the CcO molecule, designated the H pathway, located near heme *a* (20). The driving force for active proton transport is electrostatic repulsion between the proton in the hydrogen-bond network and the net positive charge of heme *a*. One of the critical sites for repulsion is the formyl group of heme *a*, which is hydrogen bonded to Arg38 of the CcO subunit I (21). In our study, resonance Raman spectroscopy revealed specific band shifts from 1,644 cm^{-1} to 1,673 cm^{-1}, which can be attributed to the vibration of the formyl group of heme *a*. This observation suggests that Higd1a binding causes structural changes, particularly around heme *a*, weakening the hydrogen bond between the formyl group and Arg38 of the CcO subunit I, thereby leading to the acceleration of proton pumping efficiency (22). Thus, both band shifts suggest that structural change occurs in the vicinity of heme *a* rather than a_3.

Following the resonance Raman analysis, we sought to determine the Higd1a-CcO binding site via simulation with the COOT software (23), using the previously reported structures of CcO (14) and Higd1a (24). From our structural analysis, CcO contains a cleft composed of relatively few protein subunits near the active centers (Fig. S10*A*). Notably, Higd1a was predicted to integrate into the cleft of CcO near heme *a* and Arg38 (Fig. S10), consistent with the

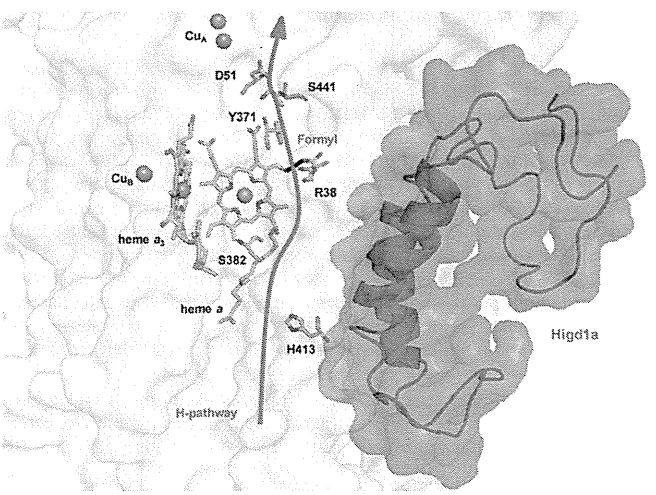

Fig. 5. Higd1a acts on the H pathway. Model depicting our docking simulation (side view) and its relationship with the H pathway. The model shows the location of Higd1a (magenta) in the CcO complex (white) and its relationship to R38 of cytochrome c oxidase subunit I and the formyl group of heme a, a component of the H pathway (red arrow).

results of the resonance Raman analysis. Thus, it is likely that Higd1a bound to the cleft of CcO, leading to swift structural change around heme *a* and Arg38 and accelerating the proton-pumping H pathway, thereby increasing CcO activity (Fig. 5). Furthermore, when we retrospectively reviewed the purification process of the CcO complex, comprising 13 subunits from the bovine heart, we found that Higd1a remained associated with CcO up to the final step, which required detergent exchange (14). This led us to speculate that Higd1a represents a 14th identified subunit of CcO that is endogenously induced by hypoxia and integrates into the open cleft of CcO to positively regulate its activity. Although the resonance Raman data and docking model simulation are consistent with the idea that Higd1a binding causes structural change around heme *a*, these data are limited because of their speculative nature. Therefore, to confirm these findings, we are currently trying to crystallize the CcO-Higd1a complex to reveal the conformational changes of CcO, particularly around the heme *a* site.

Higd1a was originally identified as a mitochondrial inner membrane protein whose expression is induced by hypoxia (25). Higd1a augments cell survival under hypoxic stress in pancreatic cells (26), and it exerts its protective effect by induction of mitochondrial fission (27). The precise relationship between these reports and our data is not clear. However, our results suggest that the elevation of CcO activity by Higd1a preserves the proton-motive force, which is prerequisite for mitochondria function, thereby leading to increased mitochondrial fission and/or the prevention of apoptosis.

The existence of a direct CcO allosteric activator suggests that there is a structural basis for the intrinsic activation in the CcO complex. To explore this idea further, a screen for small compounds that simply increase the activity of highly purified CcO in vitro has been initiated. Compounds that mimic the effect of Higd1a can preserve ATP production even under hypoxic condition, and hence are expected to exert cellular protective effects particularly when OXPHOS activity is reduced. Recent work showed that lowering the activity of OXPHOS causes the cellular senescence (28), diabetes mellitus (29), and neurodegenerative diseases (30). In addition, several currently intractable mitochondrial diseases are caused by mutations in mitochondrial genes or nuclear genes that lead to dysfunction in mitochondrial OXPHOS. Notably, decreased CcO activity is most frequently observed among patients with mitochondrial diseases (31). Therefore, small compounds that mimic the effect of Higd1a will have therapeutic potential for various acute and chronic diseases including ischemic, metabolic, and mitochondrial diseases.

Materials and Methods

Purification of Recombinant Higd1a Protein. The full-length bovine *Higd1a* cDNA was purchased from GE Healthcare. Then the coding sequence of bovine *Higd1a* was cloned in-frame with an ATG start codon, in the pET21a expression vector (Novagen for overexpression in *E. coli*). A MBP was fused in-frame at the amino terminus for purification. The resulting plasmid was transformed into BL21-Star (DE3; Invitrogen), and the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside caused the expression of MBP-Higd1a protein. The cells were sonicated and solubilized by 1% *n*-decyl-β-D-maltoside (DM). The recombinant protein was purified with amylose resin (New England Biolabs), and eluted by 20 mM maltose (pH 6.8 or 8.0, 100 mM sodium phosphate buffer containing 0.2% DM). The eluted protein was concentrated and maltose removed using Amicon Ultra-0.5 10K (Millipore).

Resonance Raman Spectroscopy. Absorption spectra of the samples were measured by a spectrophotometer (Hitachi, U3310) with the path length of 2 mm in 100 mM sodium phosphate buffer (pH 8.0) containing 0.2% DM. The reaction mixture was measured immediately and spectra were recorded every 5 min for 30 min. The protein concentration was 8 μM.

Raman scattering of the samples were measured in a cylindrical spinning cell with excitation at 413.1 nm with a Kr^+ laser (Spectra Physics, model 2060), and the incident power was 500 μW. The detector was a liquid N_2-cooled CCD detector (Roper Scientific, Spec-10: 400B/LN). Raman shifts were calibrated with indene as the frequency standard. Raman spectrum was divided by the "white light" spectrum that was determined by measuring the scattered radiation of an incandescent lamp by a white paper to compensate for the sensitivity difference of each CCD pixel and transmission curve of the notch filter to reject Rayleigh scattering. The accuracy of the peak position of well-defined Raman bands was ±1 cm^{-1}. The protein concentration was 20 μM, and the reaction mixture was incubated for 30 min, before Raman measurements.

Measurement of CcO Activity. CcO activity was measured spectrophotometrically (Shimazu, UV-2450) using a cytochrome *c* oxidase activity kit (Biochain). A total of 25 μg of mitochondrial pellets from cardiomyocytes was lysed with 1% *n*-dodecyl-β-D-maltoside (DDM), and subjected to measurement according to the manufacturer's instructions (32). Concentrations of reduced/oxidized cytochrome *c* were determined using the extinction coefficient at 550 nm of 21.84 $mM^{-1}cm^{-1}$. For in vitro measurement, cytochrome *c* (Sigma) was reduced by ascorbic acid (Wako). Recombinant MBP-Higd1a (20 μM) and hpCcO (20 μM) were incubated at 25 °C for 30 min in the presence of 0.2% DM. After incubation, the mixture and reduced cytochrome *c* were added into the assay buffer, then subjected to measurement at 30 °C (Agilent Technologies, cary300). Slopes of OD_{550} for 1 min were calculated and corrected by a value of hpCcO.

FRET-Based Measurement of Mitochondrial ATP Concentration. FRET-based measurement of mitochondrial ATP concentration in cardiomyocytes was measured as previously described (8, 33). Briefly, FRET signal was measured in cardiomyocytes infected with adenovirus encoding mit-AT1.03 with an Olympus IX-81 inverted fluorescence microscope (Olympus) using a PL APO 60×, 1.35 N.A., oil immersion objective lens (Olympus). Fluorescence emission from Mit-ATeam was imaged by using a dual cooled CCD camera (ORCA-D2; Hamamatsu Photonics) with a dichroic mirror (510 nm) and two emission filters (483/32 nm for CFP and 542/27 nm for YFP; A11400-03; Hamamatsu Photonics). Cells were illuminated using the CoolLED pE-1 excitation system (CoolLED) with a wavelength of 425 nm. Image analysis was performed using MetaMorph (Molecular Devices). The YFP/CFP emission ratio was calculated by dividing pixel by pixel (a YFP image with a CFP image after background subtraction).

Statistical Analyses. The comparison between two groups was made by *t* test (two tailed). For MASC assay, comparison was made by repeated two-way ANOVA. A value of $P < 0.05$ was considered statistically significant. Data represent mean ± SEM.

Further methods are found in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank T. Miyazaki (Cyclex) for making antibodies; Y. Okazaki and Y. Tokuzawa (Saitama Medical University) for measurement of CcO activity by cary300; Dr. Steven Coppen for critical reading of the manuscript; S. Ikezawa, E. Takada, and H. Shingu for technical assistance; M. Kobayashi, R. Maki, and the Center for Research Education in Osaka University for MS analysis; Y. Okada for secretarial support; and H. Shimada for discussion and advice. This research was supported by the Japan Society for the Promotion of Science through the "Funding Program for Next Generation World-Leading Researchers (NEXT Program)," initiated by the Council for Science and Technology Policy; grants-in-aid from the Ministry of Health, Labor, and Welfare-Japan; grants-in-aid from the Ministry of Education, Culture, Sports, Science, and Technology-Japan; and grants-in-aid from the Japan Society for the Promotion of Science. This research was also supported by grants from Takeda Science Foundation, Japan Heart Foundation, Japan Cardiovascular Research Foundation, Japan Intractable Diseases Research Foundation, Japan Foundation of Applied Enzymology, Japan Medical Association, Uehara Memorial Foundation, Mochida Memorial Foundation, Banyu Foundation, Naito Foundation, Inoue Foundation for Science, Osaka Medical Research foundation for intractable diseases, Ichiro Kanehara Foundation, and Showa Houkoukai.

1. Tsukihara T, et al. (1996) The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 A. *Science* 272(5265):1136–1144.
2. Morgan JE, Vakkasoglu AS, Lanyi JK, Gennis RB, Maeda A (2010) Coordinating the structural rearrangements associated with unidirectional proton transfer in the bacteriorhodopsin photocycle induced by deprotonation of the proton-release group: A time-resolved difference FTIR spectroscopic study. *Biochemistry* 49(15):3273–3281.
3. Aoyama H, et al. (2009) A peroxide bridge between Fe and Cu ions in the O2 reduction site of fully oxidized cytochrome c oxidase could suppress the proton pump. *Proc Natl Acad Sci USA* 106(7):2165–2169.
4. Ogura T, Kitagawa T (2004) Resonance Raman characterization of the P intermediate in the reaction of bovine cytochrome c oxidase. *Biochim Biophys Acta* 1655(1-3): 290–297.

5. Yoshikawa S, et al. (1998) Redox-coupled crystal structural changes in bovine heart cytochrome c oxidase. *Science* 280(5370):1723–1729.
6. Gospodarowicz D, Abraham JA, Schilling J (1989) Isolation and characterization of a vascular endothelial cell mitogen produced by pituitary-derived folliculo stellate cells. *Proc Natl Acad Sci USA* 86(19):7311–7315.
7. Wang GL, Semenza GL (1993) General involvement of hypoxia-inducible factor 1 in transcriptional response to hypoxia. *Proc Natl Acad Sci USA* 90(9):4304–4308.
8. Kioka H, et al. (2014) Evaluation of intramitochondrial ATP levels identifies G0/G1 switch gene 2 as a positive regulator of oxidative phosphorylation. *Proc Natl Acad Sci USA* 111(1):273–278.
9. Vukotic M, et al. (2012) Rcf1 mediates cytochrome oxidase assembly and respirasome formation, revealing heterogeneity of the enzyme complex. *Cell Metab* 15(3):336–347.
10. Strogolova V, Furness A, Robb-McGrath M, Garlich J, Stuart RA (2012) Rcf1 and Rcf2, members of the hypoxia-induced gene 1 protein family, are critical components of the mitochondrial cytochrome bc1-cytochrome c oxidase supercomplex. *Mol Cell Biol* 32(8):1363–1373.
11. Chen YC, et al. (2012) Identification of a protein mediating respiratory supercomplex stability. *Cell Metab* 15(3):348–360.
12. Fukuda R, et al. (2007) HIF-1 regulates cytochrome oxidase subunits to optimize efficiency of respiration in hypoxic cells. *Cell* 129(1):111–122.
13. Pagliarini DJ, et al. (2008) A mitochondrial protein compendium elucidates complex I disease biology. *Cell* 134(1):112–123.
14. Tsukihara T, et al. (1995) Structures of metal sites of oxidized bovine heart cytochrome c oxidase at 2.8 A. *Science* 269(5227):1069–1074.
15. Wilson DF, Gilmour MV (1967) The low-temperature spectral properties of mammalian cytochrome oxidase. I. The enzyme in intact rat-liver mitochondria. *Biochim Biophys Acta* 143(1):52–61.
16. Heibel GE, Anzenbacher P, Hildebrandt P, Schäfer G (1993) Unusual heme structure in cytochrome aa3 from Sulfolobus acidocaldarius: A resonance Raman investigation. *Biochemistry* 32(40):10878–10884.
17. Babcock GT, Callahan PM (1983) Redox-linked hydrogen bond strength changes in cytochrome a: Implications for a cytochrome oxidase proton pump. *Biochemistry* 22(10):2314–2319.
18. Fujikawa M, Yoshida M (2010) A sensitive, simple assay of mitochondrial ATP synthesis of cultured mammalian cells suitable for high-throughput analysis. *Biochem Biophys Res Commun* 401(4):538–543.
19. Imamura H, et al. (2009) Visualization of ATP levels inside single living cells with fluorescence resonance energy transfer-based genetically encoded indicators. *Proc Natl Acad Sci USA* 106(37):15651–15656.
20. Muramoto K, et al. (2010) Bovine cytochrome c oxidase structures enable O2 reduction with minimization of reactive oxygens and provide a proton-pumping gate. *Proc Natl Acad Sci USA* 107(17):7740–7745.
21. Yoshikawa S, Tsukihara T, Shinzawa-Itoh K (1996) [Crystal structure of fully oxidized cytochrome c-oxidase from the bovine heart at 2.8 A resolution]. *Biokhimiia* 61(11): 1931–1940.
22. Tsukihara T, et al. (2003) The low-spin heme of cytochrome c oxidase as the driving element of the proton-pumping process. *Proc Natl Acad Sci USA* 100(26):15304–15309.
23. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
24. Klammt C, et al. (2012) Facile backbone structure determination of human membrane proteins by NMR spectroscopy. *Nat Methods* 9(8):834–839.
25. Denko N, et al. (2000) Epigenetic regulation of gene expression in cervical cancer cells by the tumor microenvironment. *Clin Cancer Res* 6(2):480–487.
26. Wang J, et al. (2006) Pancreatic beta cells lack a low glucose and O2-inducible mitochondrial protein that augments cell survival. *Proc Natl Acad Sci USA* 103(28): 10636–10641.
27. An HJ, et al. (2013) Higd-1a interacts with Opa1 and is required for the morphological and functional integrity of mitochondria. *Proc Natl Acad Sci USA* 110(32):13014–13019.
28. Horan MP, Pichaud N, Ballard JW (2012) Review: Quantifying mitochondrial dysfunction in complex diseases of aging. *J Gerontol A Biol Sci Med Sci* 67(10): 1022–1035.
29. Saxena R, et al. (2006) Comprehensive association testing of common mitochondrial DNA variation in metabolic disease. *Am J Hum Genet* 79(1):54–61.
30. Lin MT, Beal MF (2006) Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. *Nature* 443(7113):787–795.
31. Diaz F (2010) Cytochrome c oxidase deficiency: Patients and animal models. *Biochim Biophys Acta* 1802(1):100–110.
32. Berry EA, Trumpower BL (1987) Simultaneous determination of hemes a, b, and c from pyridine hemochrome spectra. *Anal Biochem* 161(1):1–15.
33. Shintani Y, et al. (2014) Toll-like receptor 9 protects non-immune cells from stress by modulating mitochondrial ATP synthesis through the inhibition of SERCA2. *EMBO Rep* 15(4):438–445.

nature COMMUNICATIONS

ARTICLE

Received 9 Dec 2013 | Accepted 8 Jan 2015 | Published 26 Feb 2015

DOI: 10.1038/ncomms7241 OPEN

Complement C1q-induced activation of β-catenin signalling causes hypertensive arterial remodelling

Tomokazu Sumida[1,2], Atsuhiko T. Naito[1,2,3], Seitaro Nomura[1], Akito Nakagawa[3], Tomoaki Higo[3], Akihito Hashimoto[3], Katsuki Okada[3], Taku Sakai[3], Masamichi Ito[1], Toshihiro Yamaguchi[1], Toru Oka[2,3], Hiroshi Akazawa[1,2], Jong-Kook Lee[2,3], Tohru Minamino[4], Stefan Offermanns[5], Tetsuo Noda[6], Marina Botto[7], Yoshio Kobayashi[8], Hiroyuki Morita[1], Ichiro Manabe[1], Toshio Nagai[8], Ichiro Shiojima[2,9] & Issei Komuro[1,2,3]

Hypertension induces structural remodelling of arteries, which leads to arteriosclerosis and end-organ damage. Hyperplasia of vascular smooth muscle cells (VSMCs) and infiltration of immune cells are the hallmark of hypertensive arterial remodelling. However, the precise molecular mechanisms of arterial remodelling remain elusive. We have recently reported that complement C1q activates β-catenin signalling independent of Wnts. Here, we show a critical role of complement C1-induced activation of β-catenin signalling in hypertensive arterial remodelling. Activation of β-catenin and proliferation of VSMCs were observed after blood-pressure elevation, which were prevented by genetic and chemical inhibition of β-catenin signalling. Macrophage depletion and *C1qa* gene deletion attenuated the hypertension-induced β-catenin signalling, proliferation of VSMCs and pathological arterial remodelling. Our findings unveil the link between complement C1 and arterial remodelling and suggest that C1-induced activation of β-catenin signalling becomes a novel therapeutic target to prevent arteriosclerosis in patients with hypertension.

[1] Department of Cardiovascular Medicine, The University of Tokyo Graduate School of Medicine, Tokyo 113-8655, Japan. [2] CREST, Japan Science and Technology Agency, Tokyo 102-0075, Japan. [3] Department of Cardiovascular Medicine, Osaka University Graduate School of Medicine, Osaka 565-0871, Japan. [4] Department of Cardiovascular Biology and Medicine, Niigata University Graduate School of Medical and Dental Sciences, Niigata 951-8510, Japan. [5] Department of Pharmacology, Max-Planck-Institute for Heart and Lung Research, Bad Nauheim D-61231, Germany. [6] Department of Cell Biology, The Cancer Institute, Japanese Foundation for Cancer Research, Tokyo 135-8550, Japan. [7] Centre for Complement and Inflammation Research, Department of Medicine, Imperial College London, London SW7 2AZ, UK. [8] Department of Cardiovascular Medicine, Chiba University Graduate School of Medicine, Chiba 260-8670, Japan. [9] Department of Medicine II, Kansai Medical University, Osaka 573-1191, Japan. Correspondence and requests for materials should be addressed to I.K. (email: komuro-tky@umin.ac.jp).

Hypertension, as the leading risk factor for various cardiovascular diseases, caused 9.4 million deaths in 2010 (ref. 1). In 2000, nearly one billion people have hypertension worldwide, and that number is estimated to increase to 1.5 billion by 2025 (ref. 2). One direct physiological consequence of blood-pressure elevation is the structural remodelling of the arteries. Prolonged high blood pressure causes arterial degeneration due to lack of capability of optimized remodelling, leading to pathological conditions such as arteriosclerosis and end-organ damage[3,4]. Hypertrophy/hyperplasia of vascular smooth muscle cells (VSMCs) and infiltration of inflammatory cells are the characteristics of pathological arterial remodelling[4–6]. A variety of humoral factors such as growth factors, proteases and cytokines, secreted by infiltrated immune cells, have been reported to be involved in VSMC proliferation[3,7], however, precise molecular and cellular mechanisms of how hypertensive arterial remodelling is developed remain elusive.

Wnt/β-catenin signalling is an evolutionarily conserved intracellular signalling that plays an important role in embryonic development and various diseases[7,8]. The Wnt/β-catenin pathway is the most well-understood signalling cascade initiated by Wnt proteins and acts as a mitogenic signal during the development of multiple organs, and the aberrant activation of Wnt/β-catenin signalling pathway is often associated with cancer[9–11]. Wnt/β-catenin signalling regulates the proliferation and differentiation of smooth muscle cells during embryonic and postnatal angiogenesis[12,13]. Furthermore, activation of Wnt/β-catenin signalling is implicated in VSMC proliferation during intimal thickening after vascular injury[14,15].

Recently, we have reported that complement protein C1q, an initiator of the classical complement pathway, activates canonical Wnt signalling in a complement cascade-independent manner[16]. Given that the major source of complement C1q is monocyte-derived cells[17] and that macrophages (Mϕs) within the aortic wall play a vital role in pathogenesis of arterial remodelling, we hypothesized that aortic Mϕ-derived C1q activates β-catenin signalling and induces proliferation of VSMCs. In the present study, we elucidated that complement C1q, which is mainly secreted by alternatively activated aortic Mϕs, is involved in hypertensive arterial remodelling via activation of β-catenin signalling.

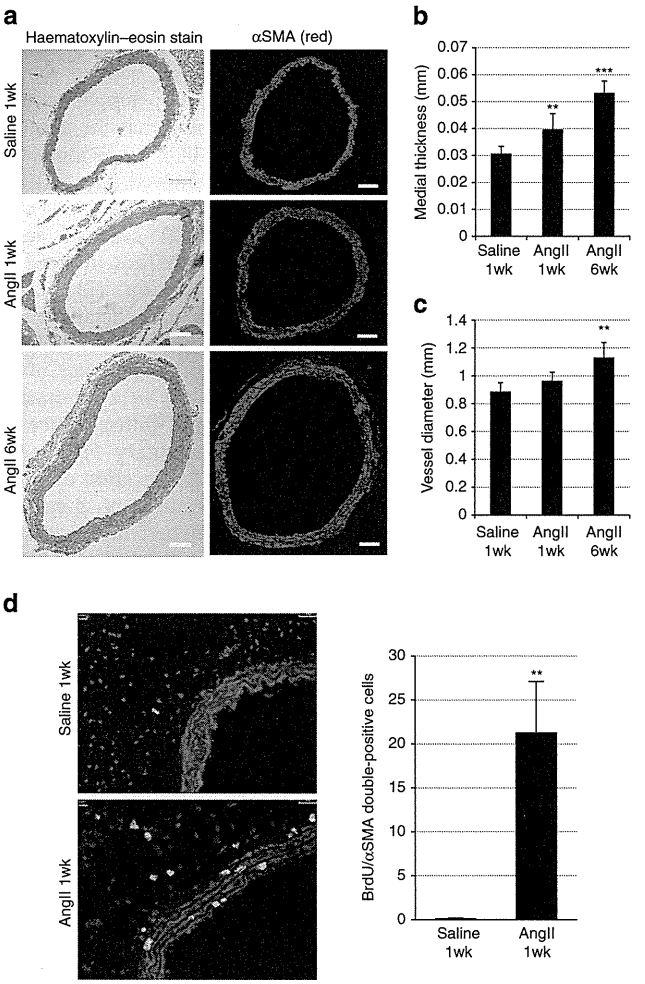

Figure 1 | Proliferation of VSMCs is observed at the initial stage of AngII-induced arterial remodelling. (**a**) Haematoxylin and eosin staining and immunostaining for α-smooth muscle actin (αSMA) of the abdominal aorta from 1-week saline-infused mice (saline 1wk), 1-week AngII-infused mice (AngII 1wk) and 6-week AngII-infused mice (AngII 6wk) ($n=5$–7). Scale bar, 100 μm. (**b**,**c**) Morphometric analysis. (**b**) Medial thickness and (**c**) vessel diameter were calculated using ImageJ. $^{**}P<0.01$, $^{***}P<0.001$ versus saline-infused mice ($n=5$–9). (**d**) Aortic tissues were immunostained for BrdU (green) and αSMA (red). Scale bar, 100 μm. The number of double-positive (BrdU(+)/αSMA(+)) cells per section is shown. $^{**}P<0.01$ versus saline-infused mice ($n=5$–7). Statistical significance was determined using one-way analysis of variance with Turkey's *post hoc* test for **b** and **c**, and the unpaired two-tailed Student's *t*-test for **d**. Results are represented as mean ± s.d.

Results

VSMC proliferation as an early event in hypertension. Infusion of angiotensin II (AngII) ($1.8\,\mu g\,kg^{-1}\,min^{-1}$) to male C57BL/6 mice increased blood pressure by ~70 mm Hg and promoted arterial remodelling characterized by thickening and dilation of the abdominal aorta at 6 weeks after AngII infusion (Fig. 1a–c), as reported previously[18]. The number of 5-bromo-2′deoxyuridine (BrdU)-positive, proliferating VSMCs was significantly increased as early as 1 week after AngII infusion (Fig. 1d), when the gross structural remodelling of the arteries was not observed. These results suggest that VSMC proliferation is one of the earliest events that occur in response to AngII-induced blood-pressure elevation.

β-catenin signalling is activated at early hypertension. Various growth factors and G-protein-coupled receptor agonists induce VSMC proliferation[7,19], and their mitogenic effects are often given by the activation of the mitogen-activated protein kinases, especially the extracellular signal regulated kinase (ERK) signalling pathway[5,20]. Therefore, we first examined whether ERK signalling was activated concurrently with VSMC proliferation after AngII infusion, and found that there was no increase in the phosphorylation levels of ERK1/2 (Supplementary Fig. 1a). Moreover, AngII activated ERK1/2 but had no effect on proliferation of cultured human aortic smooth muscle cells (HASMCs) (Supplementary Fig. 1b,c), suggesting that signalling pathways other than the ERK pathway play a key role in VSMC proliferation at the early stage of AngII-induced hypertension.

Wnt/β-catenin signalling has been reported to regulate proliferation and differentiation of smooth muscle cells during embryonic and postnatal angiogenesis[12,13] and intimal thickening after vascular injury[14,15]. We therefore tested whether β-catenin signalling was involved in VSMC proliferation after AngII infusion. Activation of β-catenin signalling upregulates the expression of Wnt/β-catenin target genes including *Axin2*

(ref. 10). At 1 week after AngII infusion, the number of LacZ-positive VSMCs was increased in Wnt reporter Axin2^{LacZ} mice[11] (Fig. 2a) and the aortic VSMCs were strongly stained for β-catenin and Axin2 in wild-type mice (Fig. 2b; Supplementary Fig. 2a,b). Western blot (WB) analysis revealed that the amount of non-phosphorylated, active β-catenin (ABC)[21] was increased (Fig. 2c), and gene expression analysis also showed upregulation of Wnt/β-catenin target genes in aortic tissue (Fig. 2d). Although β-catenin was also expressed in endothelial cells of aorta (Supplementary Fig. 2c), Axin2 expression was not induced in endothelial cells 1 week after AngII infusion (Fig. 2a,b; Supplementary Fig. 2b), suggesting that β-catenin signalling was more potently activated in aortic VSMCs rather than endothelial cells at the early stage of hypertension. Both activation of β-catenin signalling and proliferation of VSMCs were attenuated when we normalized blood pressure in AngII-infused mice by oral administration of hydralazine (Supplementary Fig. 3a–c). Taken together, these results suggested that the elevation of blood pressure was responsible in part for activation of β-catenin signalling in the aortic VSMCs during hypertensive arterial remodelling.

β-catenin signal activation induces VSMC proliferation. To determine whether the activation of β-catenin signalling induces proliferation of VSMCs, we first treated cultured HASMCs with Wnt3A, a canonical Wnt ligand, or lithium chloride, which stabilizes β-catenin by inhibiting glycogen synthase kinase-3. Both Wnt3A and lithium chloride treatment activated β-catenin signalling and induced proliferation of HASMCs (Fig. 3a,b). Moreover, overexpression of constitutively ABC also induced proliferation of HASMCs (Fig. 3c), suggesting that activation of β-catenin signalling is sufficient to induce VSMC proliferation *in vitro*. On the other hand, AngII treatment did not induce activation of β-catenin signalling or proliferation of HASMCs (Supplementary Figs 1c and 3d), but induced cellular hypertrophy of HASMCs (Supplementary Fig. 3e). We next examined whether activation of β-catenin signalling was required for AngII-induced VSMC proliferation *in vivo*. Intraperitoneal injection of PKF115-584, a small-molecule inhibitor of β-catenin signalling[22], suppressed VSMC proliferation without lowering blood pressure at 1 week after AngII infusion (Fig. 3d; Supplementary Fig. 3f). To further determine the role of the β-catenin signal in VSMCs, we generated smooth muscle cell-specific, tamoxifen-inducible, β-catenin knockout mice by crossing SMMHC-CreER^{T2} mice[23] with *Ctnnb1*$^{flox/flox}$ mice (SMMHC/β-catenin CKO). The expression levels of β-catenin and Axin2 in aortic tissue were downregulated when they were treated with tamoxifen (Fig. 4a–c). There was no change in systolic blood pressure of SMMHC/β-catenin CKO mice as compared with control mice (Fig. 4d). TdT-mediated dUTP nick end labelling staining revealed that there was no change in cell death and cell density in aortic media after tamoxifen treatment (Fig. 4e,f). There was less proliferation of VSMCs after AngII infusion in SMMHC/β-catenin CKO mice as compared with control mice (Fig. 4g), suggesting that activation of β-catenin signalling substantially contributes to enhanced proliferation of VSMCs at the early phase of hypertensive arterial remodelling.

Recruitment of Mϕs is essential for β-catenin signalling. Mϕs were recruited to the aortic adventitia by 1 week after AngII

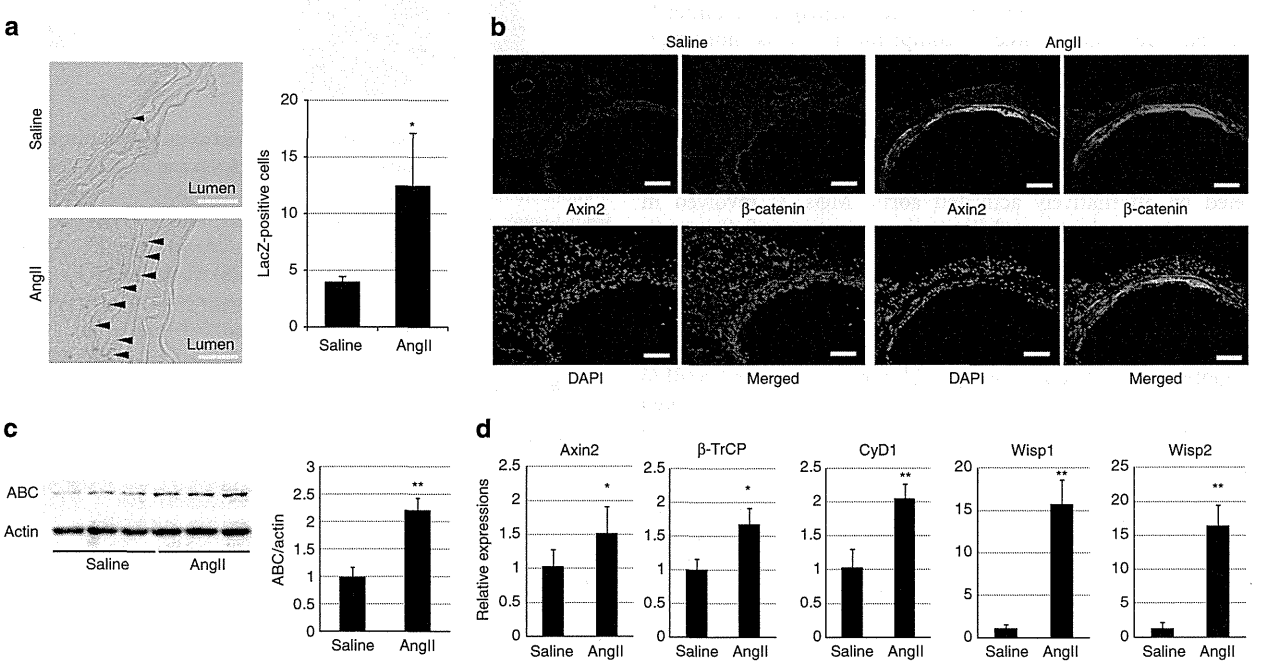

Figure 2 | β-catenin signalling is activated in the aortic media at the early stage of hypertension. (**a**) β-galactosidase staining of the aortic tissue from 1-week saline- or AngII-infused Axin2^{LacZ} mice. Arrowheads indicate β-galactosidase-positive nuclei. Scale bar, 50 μm. The number of LacZ-positive cells in the aortic media from Axin2^{LacZ} mice is shown. *$P<0.05$ versus saline-infused mice ($n=5$). (**b**) Aortic tissues from 1-week saline- or AngII-infused mice were immunostained for Axin2 (green) and β-catenin (red). Scale bar, 100 μm. (**c**) Western blot analysis for non-phosphorylated active β-catenin (ABC) in the aortic tissues from 1-week saline- or AngII-infused mice. Activation of β-catenin signalling was quantified by measuring the relative level of ABC over actin. The values are shown as fold induction over saline-infused mice. **$P<0.01$ versus saline-infused mice ($n=6$, 7). (**d**) Real-time PCR analysis for the expression levels of the β-catenin target genes (*Axin2, β-TrCP, cyclin D1 (CyD1), Wisp1* and *Wisp2*) in the aortic tissue from 1-week saline- or AngII-infused mice. The values are shown as fold induction over saline-infused mice. *$P<0.05$, **$P<0.01$ versus saline ($n=6$). Statistical significance was determined using the unpaired two-tailed Mann–Whitney *U*-test for **a**, **c** and **d**. Results are represented as mean ± s.d. DAPI, 4′,6-diamidino-2-phenylindole.

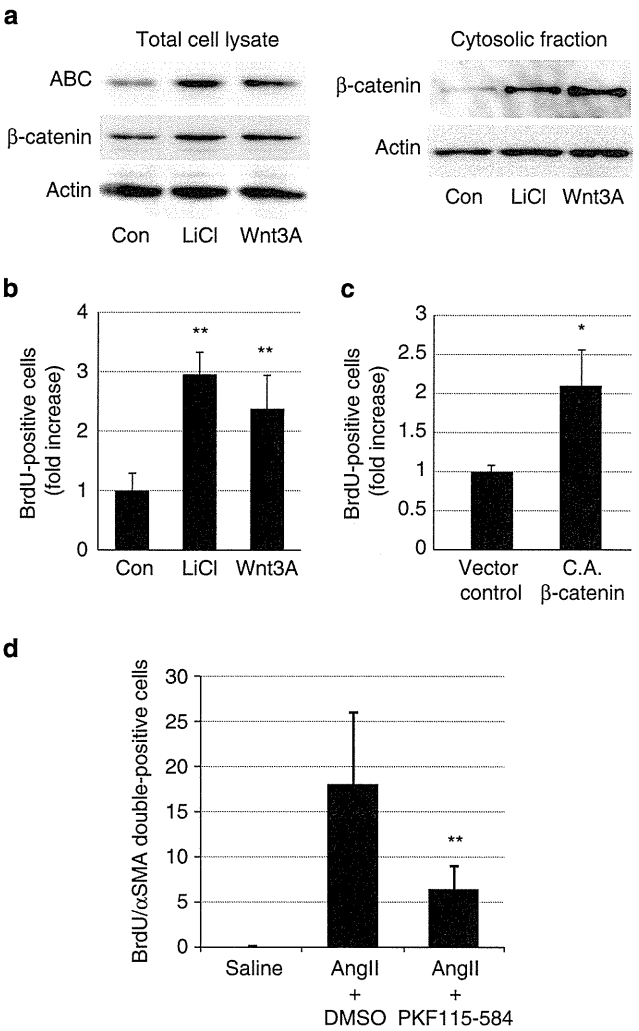

Figure 3 | Activation of β-catenin signalling induces VSMC proliferation. (**a**) Western blot analysis. HASMCs were treated with LiCl (10 mM) or Wnt3A (80 ng ml^{-1}), and the amount of ABC and cytosolic β-catenin was analysed. Activation of β-catenin signalling was quantified by measuring the relative level of ABC over actin. (**b**) The number of BrdU(+) HASMCs after LiCl (10 mM) or Wnt3A (80 ng ml^{-1}) treatment. The values are shown as fold induction over non-treated HASMCs (Con). $^{**}P<0.01$ versus non-treated HASMCs ($n=4$). (**c**) The number of BrdU(+) HASMCs after infection with control retrovirus (vector control) or with constitutively active β-catenin (CA β-catenin) whose phosphorylation sites at the N terminus are all mutated. The values are shown as fold induction over control vector-transfected HASMCs. $^{*}P<0.05$ versus control vector-transfected HASMCs ($n=3$). (**d**) The number of double-positive (BrdU(+)/αSMA(+)) cells per aortic section from saline-infused mice, AngII-infused mice treated with DMSO (solvent), and AngII-infused mice treated with PKF115-584. $^{**}P<0.01$ versus AngII-infused mice treated with DMSO ($n=8$). Statistical significance was determined using one-way analysis of variance with Turkey's *post hoc* test for **b**, the unpaired two-tailed Student's *t*-test for **c** and the Kruskal–Wallis test with Dunn's correction for multiple comparisons for **d**. Results are represented as mean ± s.d.

infusion (Fig. 5a; Supplementary Fig. 4a). To elucidate the role of these Mϕs, we injected liposomes containing clodronate (Clo-Lip)[24] before AngII infusion to deplete the cells of monocyte lineage. Clo-Lip treatment reduced the number of infiltrating Mϕs (Fig. 5b; Supplementary Fig. 4a), and strongly suppressed activation of β-catenin signalling and proliferation of VSMCs at 1 week after AngII infusion (Fig. 5c–e). These results suggest that infiltrating Mϕs play critical roles in activation of β-catenin signalling and proliferation of VSMCs at the early phase of arterial remodelling.

M2 Mϕs secrete a potent activator of β-catenin signalling. Mϕs change their phenotypes in response to various environmental factors. Classically activated Mϕs or M1 Mϕs are induced by lipopolysaccharide (LPS), interferon-γ and tumour necrosis factor, and secrete pro-inflammatory cytokines to increase their killing ability. On the other hand, alternatively activated Mϕs or M2 Mϕs are induced by interleukin (IL)-4 and IL-13, and contribute to tissue remodelling and wound healing[25]. To determine the phenotype of Mϕs that accumulate into the aortic wall at the early stage of arterial remodelling, we used flow cytometric analysis to characterize M1-type Mϕ for Ly6c expression and M2-type Mϕ for CD206 expression[26]. Approximately 40% of the aortic Mϕs from AngII-infused mice expressed CD206 (CD11b + F4/80 + CD206 + Ly6c −: M2 type) and ∼10% of them expressed Ly6c (CD11b + F4/80 + CD206-Ly6c +: M1 type) (Fig. 6a; Supplementary Fig. 4b). Immunofluorescent analysis also showed large numbers of CD206-positive Mϕs in aortic adventitia (Fig. 6b). These results prompted us to postulate that M2-type Mϕs secrete humoral factors that activate β-catenin signalling and induce proliferation of VSMCs. To test this hypothesis, we first examined the effects of humoral factors secreted by M2 Mϕs on VSMCs. The M1 or M2 phenotype was elicited *in vitro* by treating Raw264.7 cells with LPS or IL-4, respectively. Conditioned media (CM) from Raw264.7 cells treated with PBS (Con Raw264.7 CM), LPS (LPS Raw264.7 CM) or IL-4 (IL-4 Raw264.7 CM) were added to HASMCs. IL-4 Raw264.7 CM activated β-catenin signalling and promoted cell proliferation in HASMCs more potently than Con Raw264.7 CM or LPS Raw264.7 CM (Fig. 6c,d). As Raw264.7 cells are known to be already activated, we performed the same experiments using bone marrow-derived Mϕs (BMDMs), and found that CM from IL-4-treated BMDMs (IL-4 BMDM CM) activated β-catenin signalling and induced proliferation of VSMCs more than CM from PBS-treated BMDMs (Con BMDM CM) or LPS-treated BMDMs (LPS BMDM CM) (Supplementary Fig. 4c,d). These results collectively suggest that M2 Mϕs, recruited to the vessel wall in response to blood-pressure elevation, secrete a substance that activates β-catenin signalling in VSMCs, thereby inducing VSMC proliferation in a paracrine manner. Expression levels of canonical Wnt ligands (Wnt1, 2, 2b,3, 3a, 7a, 7b, 8a, 8b, 10a and 10b), which could activate the β-catenin signalling pathway[10,27–29], were not upregulated in Raw264.7 cells or BMDMs by IL-4-induced M2 polarization (Supplementary Fig. 4e,f).

Complement C1q is a Mϕ-derived β-catenin signal activator. We have recently reported that complement protein C1q, an initiator of the classical complement pathway, activates β-catenin signalling and induces aging-associated impairment of skeletal muscle regeneration[16]. Given that monocyte/Mϕ lineage has been shown to be the major source of C1q biosynthesis[17] and tumour-associated Mϕs have been reported to produce complement C1q[30], we examined whether β-catenin signalling was activated in VSMCs by M2 Mϕ-derived C1q. *In vitro*, M2-polarized Raw264.7 Mϕs by IL-4 treatment markedly increased *C1qa* gene expression compared with PBS-treated or LPS-treated Raw264.7 Mϕs (Fig. 7a). To examine whether C1q is predominantly produced by M2 Mϕs rather than M0 or M1 Mϕs *in vivo*, we sorted M0-type (CD11b + F4/80 +

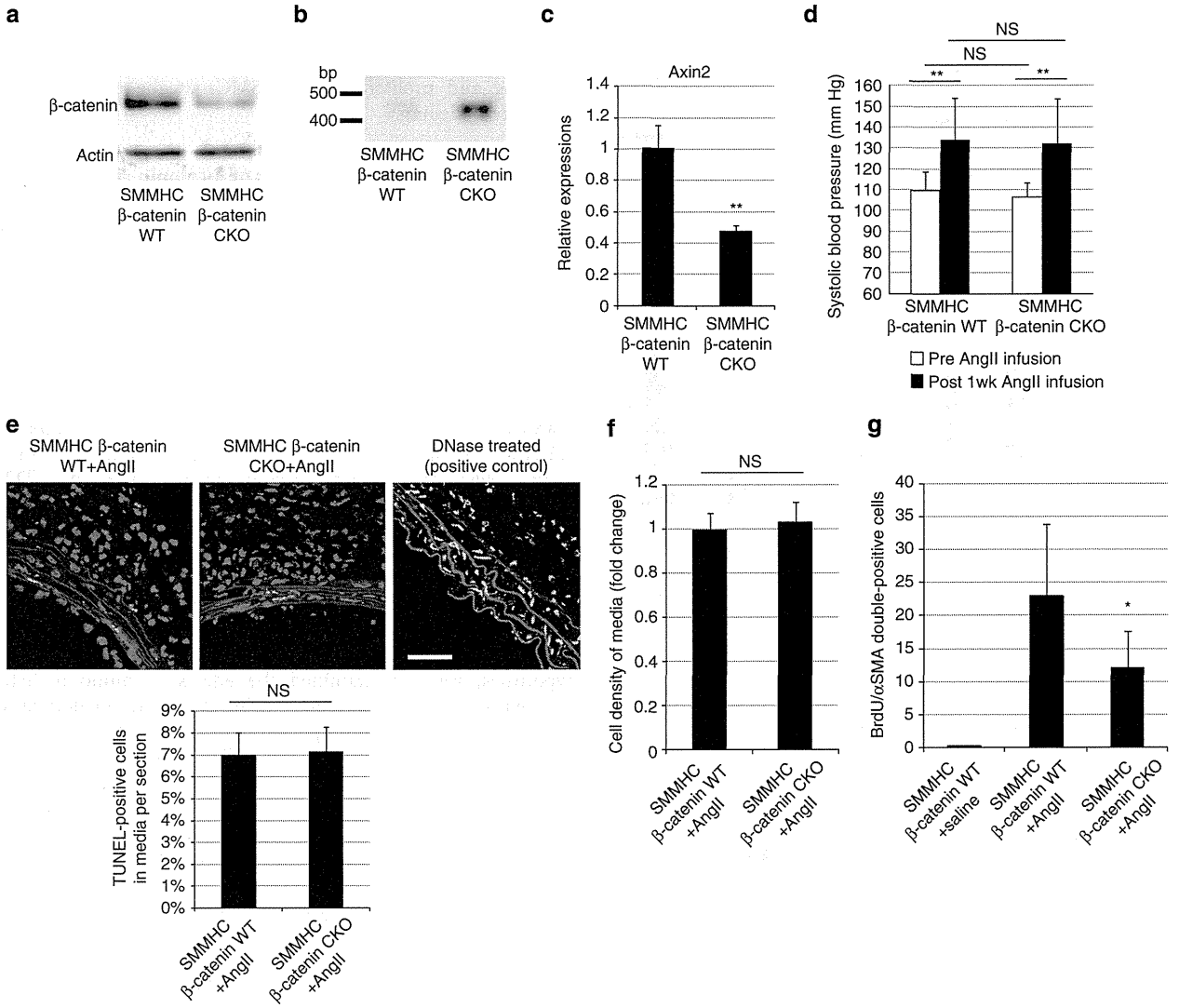

Figure 4 | β-catenin signal activation is responsible for VSMC proliferation after AngII infusion. (**a**) Western blot analysis. The amounts of β-catenin in the aortic tissues from SMMHC-CreER^{T2}:*Ctnnb1*$^{+/+}$ mice (SMMHC-β-catenin wild type (WT)) and SMMHC-CreER^{T2}:*Ctnnb1*$^{flox/flox}$ mice (SMMHC-β-catenin CKO) were analysed in aortic tissues isolated 6 days after the final tamoxifen treatment. (**b**) PCR analysis of aortic tissue DNA. DNA extracted from aortic tissues of tamoxifen-treated SMMHC-β-catenin WT and SMMHC-β-catenin CKO mice were amplified with a PCR primer set designed for detecting the null allele. (**c**) Real-time PCR analysis for the expression level of the *Axin2* gene (one of the major Wnt/β-catenin target genes) in the aortic tissue isolated from tamoxifen-treated SMMHC-β-catenin WT and SMMHC-β-catenin CKO mice. The values are shown as fold induction over SMMHC-β-catenin WT mice. $^{**}P<0.01$ *versus* SMMHC-β-catenin WT mice. (**d**) Systolic blood pressure before and after AngII infusion for 1 week. There was no difference in systolic blood pressure between SMMHC-β-catenin WT mice and SMMHC-β-catenin CKO mice before and after AngII infusion. $^{**}P<0.01$ versus Post 1wk AngII infusion. (**e**) TdT-mediated dUTP nick end labelling (TUNEL) staining of aortic tissue and percentage of TUNEL-positive cells. TUNEL staining of aortic tissue from SMMHC-β-catenin WT and SMMHC-β-catenin CKO mice 1 week after AngII infusion. The DNase (TACS nuclease)-treated section is presented as a positive control. Percentage of TUNEL-positive cells per total cells in aortic media was calculated. Scale bar, 50 μm. (**f**) Cell density of aortic media was calculated by measuring the number of αSMA-positive cells per field of view size (40 × 40 μm^2). (**g**) The number of double-positive (BrdU(+)/αSMA(+)) cells per aortic section from 1-week saline-infused SMMHC-CreER^{T2}:*Ctnnb1*$^{+/+}$ mice (SMMHC-β-catenin WT +saline), 1-week AngII-infused SMMHC-CreER^{T2}:*Ctnnb1*$^{+/+}$ mice (SMMHC-β-catenin WT + AngII) and SMMHC-CreER^{T2}:*Ctnnb1*$^{flox/flox}$ mice (SMMHC-β-catenin CKO + AngII). $^{*}P<0.05$ versus SMMHC-β-catenin WT + AngII ($n=12$). The values are shown as fold induction over SMMHC-β-catenin WT mice ($n=5$). Statistical significance was determined using the unpaired two-tailed Student's *t*-test for **c**, **e** and **f**, two-way analysis of variance followed by Sidak's multiple comparisons test for **d** and the Kruskal–Wallis test with Dunn's correction for multiple comparisons for **g**. Results are represented as mean ± s.d. NS, not significant.

CD206 − Ly6c −), M1-type (CD11b + F4/80 + CD206 − Ly6c +) and M2-type (CD11b + F4/80 + CD206 + Ly6c −) Mϕs from aortic tissue and compared the expression levels of C1q among three types of Mϕs, and found that M2-type Mϕs expressed more C1q than M0- or M1-type Mϕs (Fig. 7b; Supplementary Fig. 4b).

The direct effect of C1q on HASMCs was examined *in vitro*. C1q protein activated β-catenin signalling and promoted proliferation of cultured HASMCs in a dose-dependent manner, and these effects were suppressed by C1-inhibitor (C1-INH), an endogenous inhibitor of C1r and C1s (Fig. 7c,d; Supplementary Fig. 5a). Moreover, treatment with the C1