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, Tsukhihara 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.

Trans-ancestry mutational landscape of hepatocellular carcinoma genomes

Yasushi Totoki[1,14], Kenji Tatsuno[2,14], Kyle R Covington[3,14], Hiroki Ueda[2], Chad J Creighton[3,4], Mamoru Kato[1], Shingo Tsuji[2], Lawrence A Donehower[5], Betty L Slagle[5], Hiromi Nakamura[1], Shogo Yamamoto[2], Eve Shinbrot[3], Natsuko Hama[1], Megan Lehmkuhl[3], Fumie Hosoda[1], Yasuhito Arai[1], Kim Walker[3], Mahmoud Dahdouli[3], Kengo Gotoh[2], Genta Nagae[2], Marie-Claude Gingras[3], Donna M Muzny[3], Hidenori Ojima[6], Kazuaki Shimada[7], Yutaka Midorikawa[8], John A Goss[9], Ronald Cotton[9], Akimasa Hayashi[2,10], Junji Shibahara[10], Shumpei Ishikawa[10], Jacfranz Guiteau[9], Mariko Tanaka[10], Tomoko Urushidate[1], Shoko Ohashi[1], Naoko Okada[1], Harsha Doddapaneni[3], Min Wang[3], Yiming Zhu[3], Huyen Dinh[3], Takuji Okusaka[11], Norihiro Kokudo[12], Tomoo Kosuge[7], Tadatoshi Takayama[8], Masashi Fukayama[10], Richard A Gibbs[3], David A Wheeler[3], Hiroyuki Aburatani[2] & Tatsuhiro Shibata[1,13]

Diverse epidemiological factors are associated with hepatocellular carcinoma (HCC) prevalence in different populations. However, the global landscape of the genetic changes in HCC genomes underpinning different epidemiological and ancestral backgrounds still remains uncharted. Here a collection of data from 503 liver cancer genomes from different populations uncovered 30 candidate driver genes and 11 core pathway modules. Furthermore, a collaboration of two large-scale cancer genome projects comparatively analyzed the trans-ancestry substitution signatures in 608 liver cancer cases and identified unique mutational signatures that predominantly contribute to Asian cases. This work elucidates previously unexplored ancestry-associated mutational processes in HCC development. A combination of hotspot *TERT* promoter mutation, *TERT* focal amplification and viral genome integration occurs in more than 68% of cases, implicating *TERT* as a central and ancestry-independent node of hepatocarcinogenesis. Newly identified alterations in genes encoding metabolic enzymes, chromatin remodelers and a high proportion of mTOR pathway activations offer potential therapeutic and diagnostic opportunities.

HCC is the third leading cause of cancer deaths worldwide[1,2]. Epidemiologically, the incidence of HCC shows marked variance across geographical regions and ancestry groups and between the sexes[3]. HCC incidence predominates in East Asia and Africa, and rapid increases in prevalence have occurred in Western countries[2]. Multiple etiological cofactors are associated with liver cancer, and their contributions might additionally differ according to ancestry. Hepatitis B virus (HBV) infection is dominant in East Asia and Africa, whereas hepatitis C virus (HCV) infection among HCC cases is frequent in Japan. Aflatoxin B1 exposure is a strong risk factor of HCC in China and Africa, whereas alcohol intake is a major etiological factor for HCC in Western countries[3–5]. The average male/female ratio for HCC incidence is greater than two, which could be owing to different environmental exposures or hormone levels[6]. Overlapping but partially distinctive epidemiological backgrounds, such as liver fluke infection, were associated with intrahepatic cholangiocarcinoma (IHCC), another type of liver cancer[5]. Here we conducted the first trans-ancestry HCC genome sequencing research under the umbrella of the International Cancer Genome Consortium (ICGC)[7] and The Cancer Genome Atlas (TCGA)[8]. Thus far, this study represents the largest genomic profiling of liver cancers (608 cases) and compares ancestry groups (Japanese, Asian and European) with distinctive etiological cofactors. This genome data set also uncovers an extensive landscape of driver genetic alterations in HCC.

RESULTS

Whole-exome and oncovirome sequencing of liver cancers

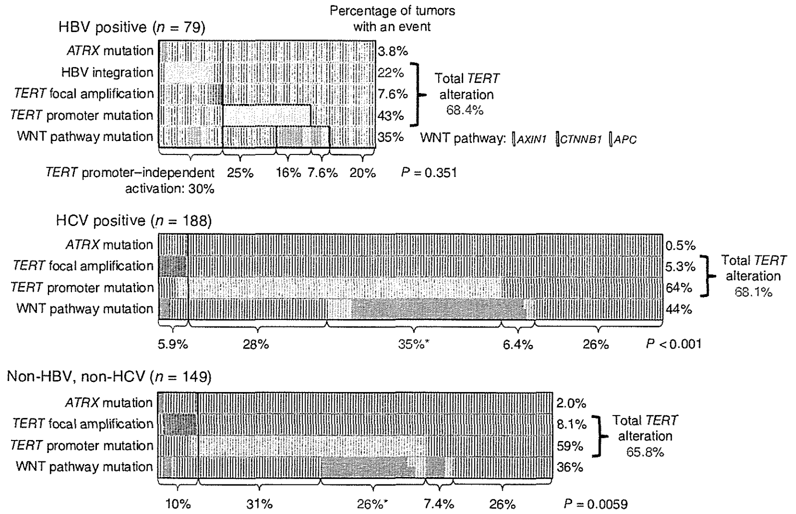
As an ICGC liver cancer project, we collected 503 pairs (413 cases in the Japanese cohort and 90 cases in the US cohort) of liver cancers (488 HCC and 15 IHCC) and matched non-cancerous liver tissues

---

[1]Division of Cancer Genomics, National Cancer Center Research Institute, Tokyo, Japan. [2]Genome Science Division, Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan. [3]Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas, USA. [4]Department of Medicine, Baylor College of Medicine, Houston, Texas, USA. [5]Department of Molecular Virology and Microbiology, Baylor College of Medicine, Houston, Texas, USA. [6]Division of Molecular Pathology, National Cancer Center Research Institute, Tokyo, Japan. [7]Hepatobiliary and Pancreatic Surgery Division, National Cancer Center Hospital, Tokyo, Japan. [8]Department of Digestive Surgery, Nihon University School of Medicine, Tokyo, Japan. [9]Department of Surgery, Baylor College of Medicine, Houston, Texas, USA. [10]Department of Pathology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. [11]Hepatobiliary and Pancreatic Oncology Division, National Cancer Center Hospital, Tokyo, Japan. [12]Hepato-Biliary-Pancreatic Surgery Division, Department of Surgery, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. [13]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan. [14]These authors contributed equally to this work. Correspondence should be addressed to D.A.W. (wheeler@bcm.edu), H.A. (haburata-tky@umin.ac.jp) or T.S. (tashibat@ncc.go.jp).

Received 31 December 2013; accepted 3 October 2014; published online 2 November 2014; doi:10.1038/ng.3126

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**).

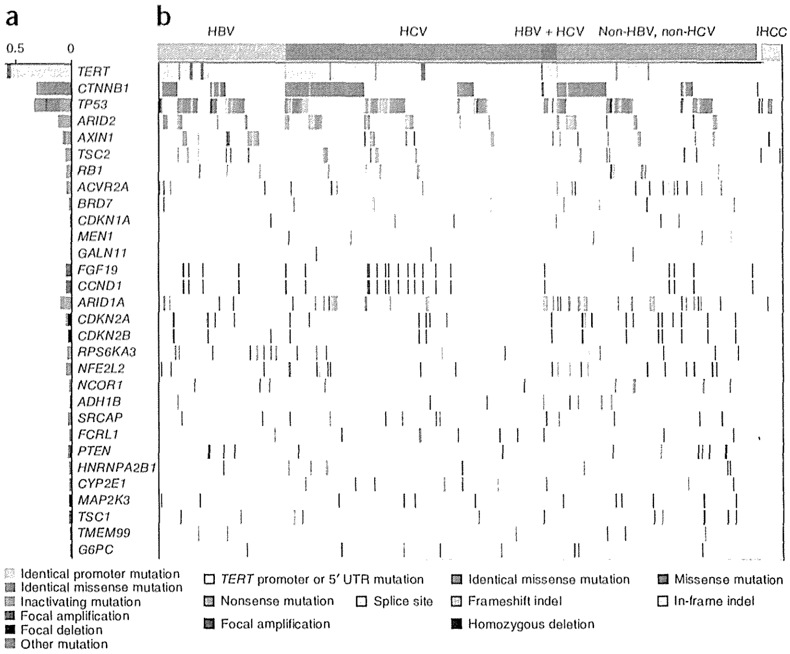


Figure 2 Significant cancer driver genes in HCC. An overview of significant driver genes in HCC. Shown are genes with statistically significant mutations or focal CNAs (**a**) and their alterations in each sample classified by the status of hepatitis virus infection (**b**). Genes were sorted by significant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

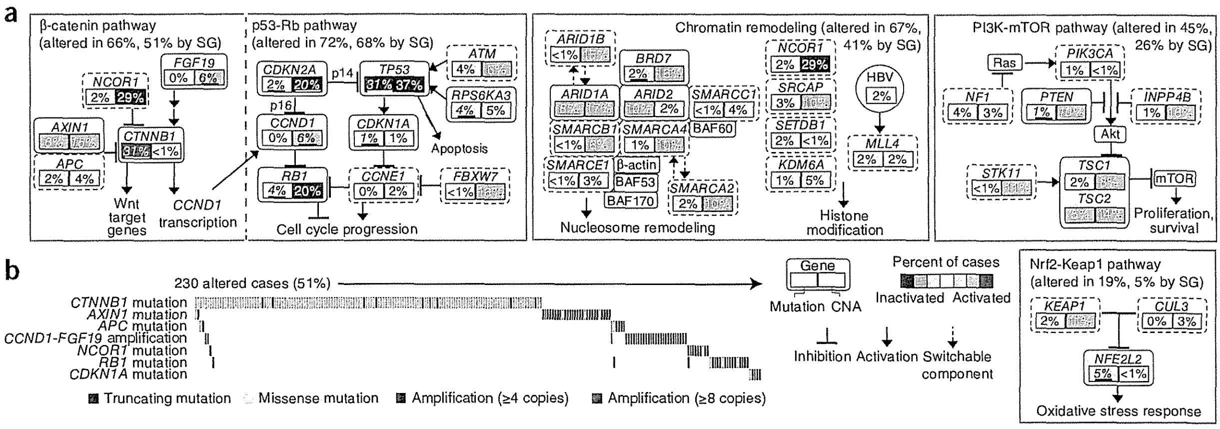
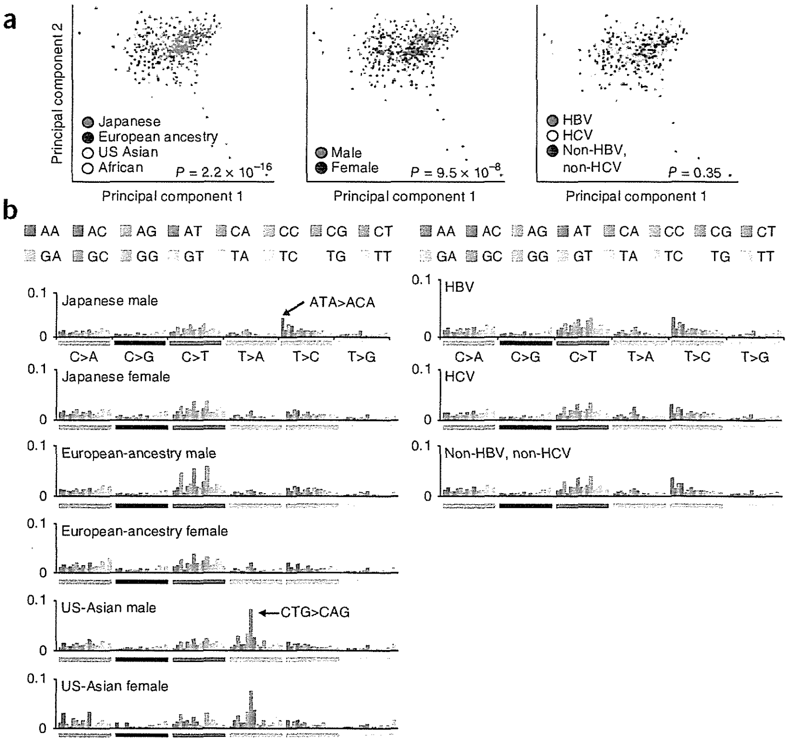


Figure 3 Oncogenic network in HCC. (**a**) Major signaling pathways involving genetic alterations in HCC. Key genes in each pathway are indicated by rectangles, with the percentages of somatic mutations and CNAs shown in the left and right portions of each rectangle, respectively. Significantly altered 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$^{\mathrm{INK4A}}$ was subject to frequent focal homozygous deletion, and the p53 target and RB regulator *CDKN1A* (encoding p21$^{\mathrm{CIP1}}$) 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

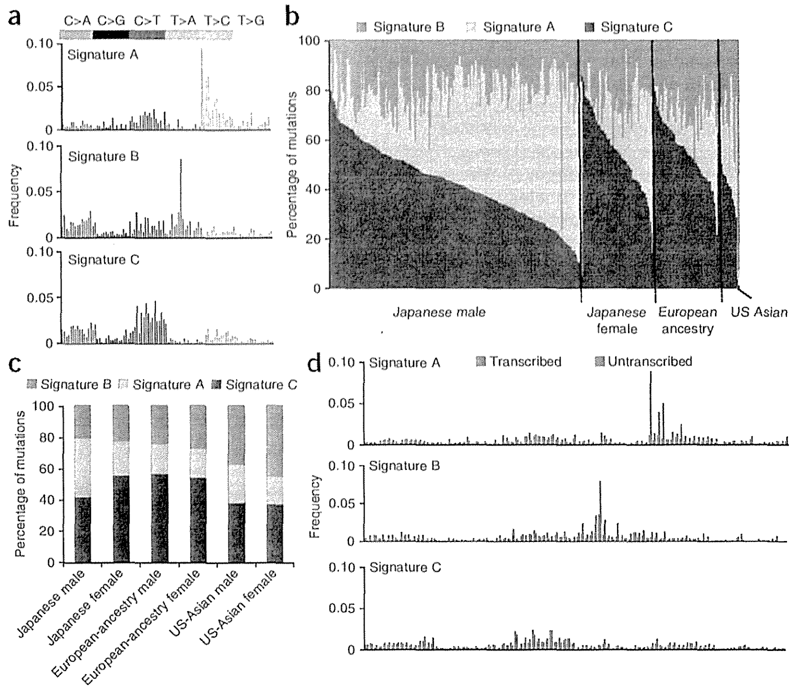


Figure 5 Ancestry-specific mutational signatures with transcriptional strand bias in the HCC genome. (**a**) The 3 mutational signatures in the HCC genome are shown according to the frequencies of 96 substitution types. The *y* axis indicates the frequency of each of the 96 substitution patterns. (**b**) 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.

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).

A novel *KCNQ1* missense mutation identified in a patient with juvenile-onset atrial fibrillation causes constitutively open I_{Ks} channels

Kanae Hasegawa, MD,[*,†] Seiko Ohno, MD, PhD,[†] Takashi Ashihara, MD, PhD,[†] Hideki Itoh, MD, PhD,[†] Wei-Guang Ding, PhD,[‡] Futoshi Toyoda, PhD,[‡] Takeru Makiyama, MD, PhD,[§] Hisaaki Aoki, MD, PhD,[¶] Yoshihide Nakamura, MD, PhD,[¶] Brian P. Delisle, PhD,[||] Hiroshi Matsuura, MD, PhD,[‡] Minoru Horie, MD, PhD[†]

From the [*]Department of Cardiovascular Biology and Medicine, Niigata University School of Medical and Dental Sciences, Niigata, Japan, [†]Department of Cardiovascular and Respiratory Medicine, Shiga University of Medical Science, Otsu, Japan, [‡]Department of Physiology, Shiga University of Medical Science, Otsu, Japan, [§]Department of Cardiovascular and Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan, [¶]Department of Pediatrics, Kinki University Faculty Medicine, Osaka, Japan, and [||]Department of Physiology, University of Kentucky, Lexington, Kentucky.

BACKGROUND Atrial fibrillation (AF) is one of the most common cardiac arrhythmias. In some patients, the disease is inheritable; however, hereditary aspects of AF remain not fully elucidated.

OBJECTIVE The purpose of this study was to identify genetic backgrounds that contribute to juvenile-onset AF and to define the mechanism.

METHODS In 30 consecutive juvenile-onset AF patients (onset age <50 years), we screened AF-related genes (*KCNQ1, KCNH2, KCNE1-3, KCNE5, KCNJ2, SCN5A*). We analyzed the function of mutant channels using whole-cell patch-clamp techniques and computer simulations.

RESULTS Among the juvenile-onset AF patients, we identified three mutations (10%): *SCN5A*-M1875T, *KCNJ2*-M301K, and *KCNQ1*-G229D. Because *KCNQ1* variant (G229D) identified in a 16-year-old boy was novel, we focused on the proband. The G229D-I_{Ks} was found to induce a large instantaneous activating component without deactivation after repolarization to −50 mV. In addition, wild-type (WT)/G229D-I_{Ks} (WT and mutant coexpression) displayed both instantaneous and time-dependent activating currents. Compared to WT-I_{Ks}, the tail current densities in WT/G229D-I_{Ks} were larger at test potentials between −130 and −40 mV but smaller at test potentials between 20 and 50 mV. Moreover, WT/G229D-I_{Ks} resulted in a negative voltage shift for current activation (−35.2 mV) and slower deactivation. WT/G229D-I_{Ks} conducted a large outward current induced by an atrial action potential waveform, and computer simulation incorporating the WT/G229D-I_{Ks} results revealed that the mutation shortened atrial but not ventricular action potential.

CONCLUSION A novel *KCNQ1*-G229D mutation identified in a juvenile-onset AF patient altered the I_{Ks} activity and kinetics, thereby increasing the arrhythmogenicity to AF.

KEYWORDS Atrial fibrillation; Juvenile-onset atrial fibrillation; Ion channel; I_{Ks}; KCNQ1

ABBREVIATIONS AF = atrial fibrillation; **AP** = action potential; **CHO** = Chinese hamster ovary; **ECG** = electrocardiogram; **QTc** = corrected QT interval; **WT** = wild type

(Heart Rhythm 2014;11:67–75) © 2014 Heart Rhythm Society. All rights reserved.

Introduction

Atrial fibrillation (AF) is the most prevalent cardiac rhythm abnormality and one of the major causes of morbidity and mortality.[1] Many risk factors predispose to AF, such as advancing age, male sex, structural heart disease, hypertension, obesity, diabetes mellitus, and hyperthyroidism.[1] In some patients, AF occurs in the absence of these risk factors; this subtype is called lone AF.[2] Genetic backgrounds have been shown to be associated with lone AF.[3] In fact, mutations in genes encoding ion channels,[4–11] gap junction proteins,[12] and signaling molecules[13] have been identified in families with AF and in isolated AF cases.[14] In 2003, Chen et al[4] first revealed among these AF-related genes the link between AF and a *KCNQ1* mutation (a gene encodes the slowly activating component of delayed rectifier K^+ current I_{Ks}). They reported a missense *KCNQ1* mutation, S140G, in familial AF, which showed a gain-of-function effect of I_{Ks}.[4,15]

This work was supported by Grants-in-Aid in Scientific Research from the Ministry of Education, Culture, Science, and Technology of Japan; a Health Sciences Research Grant from the Ministry of Health, Labour and Welfare of Japan; and Translational Research Funds from the Japan Circulation Society. **Address reprint requests and correspondence:** Dr. Minoru Horie, Department of Cardiovascular and Respiratory Medicine, Shiga University of Medical Science, Seta Tsukinowa-cho, Otsu 520-2192, Japan. E-mail address: horie@belle.shiga-med.ac.jp.

1547-5271/$-see front matter http://dx.doi.org/10.1016/j.hrthm.2013.09.073

To date, seven *KCNQ1* mutations have been reported to be associated with AF by exerting a gain-of-function effect with enhanced I_{Ks} current density with or without altered gating.[4,16–21] In addition, five of seven *KCNQ1* mutations (S140G, V141M, S209P, R231C, R231H) were identified in juvenile-onset AF patients. Among these five mutations, S140G, R231C, and R231H mutations were associated with QT prolongation. Regarding other mutations in genes that encode ion channels, functional analyses of the mutations have demonstrated either gain-of-function effects, for example, in *SCN5A* and *KCNJ2*,[5,6] or loss-of-function effects in *SCN5A* and *KCNA5*.[10,11] Intriguingly, these functional alterations are similar to those found in electrophysiologic remodeling in chronic AF.[22]

In order to clarify the genetic basis of juvenile-onset AF, we screened 30 consecutive probands for mutations in *KCNQ1, KCNH2, KCNE1-3, KCNE5, KCNJ2,* and *SCN5A*. Three heterozygous mutations were identified in *SCN5A*, *KCNJ2*, and *KCNQ1* in three probands from unrelated families (10%). We previously reported the former two mutations *SCN5A*-M1875T and *KCNJ2*-M301K.[5,6] The third missense *KCNQ1* mutation, G229D, was identified in a 16-year-old boy with AF, and it is novel. We examined the molecular mechanism underlying the *KCNQ1* mutation found in juvenile-onset AF by using a heterologous expression. We then incorporated the functional impact of the mutation into computational simulations of the atrial action potential (AP), and we found that it could contribute to shortening of the atrial AP duration leading to the arrhythmogenicity of AF.

Methods

Study subjects

The study was approved by the Institutional Ethics Committees of our institutes, and all patients provided informed consent. Thirty consecutive AF probands who developed AF at age <50 years were included in the study.

DNA isolation and genetic analysis

Genomic DNA was isolated from blood lymphocytes and screened for the entire open reading frames of *KCNQ1, KCNH2, KCNE1-3, KCNE5, KCNJ2,* and *SCN5A*. Genetic screening (except for *KCNJ2*) was performed using denaturing high-performance liquid chromatography (dHPLC WAVE System, Transgenomic, Omaha, NE, USA). Abnormal conformers and *KCNJ2* were amplified via polymerase chain reaction, and sequencing was performed using an ABI PRISM3130 DNA sequencer (Applied Biosystems, Wellesley, MA, USA). When a mutation was detected, we examined its presence in >200 Japanese healthy individuals to exclude the possibility of polymorphisms. When a mutation was detected in a proband, we checked whether or not his or her family members were also carriers.

In vitro mutagenesis

Full-length cDNA encoding human wild-type (WT) KCNQ1 (GenBank AF000571) in a pCI vector was subcloned into a pIRES2-EGFP expression vector. We engineered *KCNQ1*-G229D mutant using a site-directed mutagenesis kit (QuikChange II XL, Stratagene, La Jolla, CA, USA). The presence of mutations was confirmed by sequencing. Full-length cDNA encoding human KCNE1 (GenBank M26685) subcloned into the pCDNA3.1 expression vector was obtained by polymerase chain reaction from human heart cDNA library (Clontech Laboratories, Mountain View, CA, USA).

Electrophysiologic experiments

To assess the functional modulation by *KCNQ1* mutation, we used a heterologous expression system with the Chinese hamster ovary (CHO) cell line. Briefly, the cells were transiently transfected with *KCNQ1*-WT (0.5 μg) or *KCNQ1*-G229D (0.5 μg) or *KCNQ1*-WT (0.25 μg)/G229D (0.25 μg), and *KCNE1* (0.5 μg) plasmid DNA using Lipofectamine (Invitrogen Life Technologies, Carlsbad, CA, USA). For electrophysiologic experiments, after 48 hours of transfection, cells attached to glass coverslips were transferred to a 0.5-mL bath chamber perfused with extracellular solution at 1 to 2 mL/min. The chamber was mounted on the stage of an inverted microscope (ECLIPSE TE2000-U, Nikon, Tokyo, Japan) and maintained at 37°C. Patch-clamp experiments were conducted on GFP-positive cells.

Whole-cell membrane currents were recorded with an EPC-8 patch-clamp amplifier (HEKA Electronik, Lambrecht, Germany). Pipettes were prepared from glass capillary tube (Narishige, Tokyo, Japan) using a Sutter P-97 micropipette puller (Navato, CA, USA), and the tips were fire-polished with a microforge. Pipette resistance ranged from 2.5 to 3.5 MΩ. Pipettes were filled with a solution containing the following (in mM): 70 potassium aspartate, 40 KCl, 10 KH_2PO_4, 1 $MgSO_4$, 3 Na_2-ATP (Sigma, St. Louis, USA), 0.1 Li_2-GTP (Roche Diagnostics GmbH, Mannheim, Germany), 5 EGTA, and 5 HEPES; and pH was adjusted to 7.2 with KOH. The extracellular solution contained the following (in mM): 140 NaCl, 5.4 KCl, 1.8 $CaCl_2$, 0.5 $MgCl_2$, 0.33 NaH_2PO_4, 5.5 glucose, and 5.0 HEPES; pH was adjusted to 7.4 with NaOH. Liquid junction potential between the test solution and the pipette solution was measured to be around −10 mV and was corrected. HMR1556 (kind gift from Drs. H.J. Lang and J. Punter, Aventis Pharma Deutschland GmbH), a selective I_{Ks} blocker, was added from 10 mM stock solution in DMSO to the external solution (final DMSO concentration did not exceed 0.01%).

I_{Ks} was elicited by depolarizing voltage steps from a holding potential of −80 mV to various test potentials. I_{Ks} amplitude was determined by measuring the amplitude of the tail current elicited on repolarization to −50 mV following 2-second depolarization to 30 mV every 15 seconds and divided by the cell membrane capacitance to obtain current densities (pA/pF). I_{Ks} activation was evaluated by fitting the current–voltage relationship of the tail currents to a Boltzmann equation:

$$I_{Ktail} = 1/(1+\exp((V_h-V_m)/k))$$

where I_{Ktail} is the current amplitude density, V_h is the voltage at half-maximal activation, V_m is the test potential, and k is the slope factor. Time constants for deactivation (tau_{fast} and tau_{slow}) were obtained by fitting a two-exponential function as follows:

$$I(t) = A_{fast}\exp(-t/tau_{fast}) + A_{slow}\exp(-t/tau_{slow}) + A_0$$

where I(t) is the current amplitude at time t, A_{fast}, A_{slow}, and A_0 are constants, and tau refers to the deactivation at the tail potential.

For voltage-clamp recordings using an atrial AP waveform, we applied a waveform generated from a computer simulation of an atrial AP at 1 Hz and recorded currents at 37°C.[21] Resting membrane potential in the CHO cell was determined by current clamp after creating the whole-cell configuration.

Computer simulation

To confirm the role of the *KCNQ1*-G229D mutation, we conducted simulations of paced activation in atrial and ventricular myocytes and of paced propagation in a one-dimensional (1D) bidomain ventricular myocardial model of 9.0-mm length with transverse conductivity, mimicking a transmural section of the left ventricular free wall. Membrane kinetics of the myocytes were represented by the Courtemanche human atrial model[23] and O'Hara-Rudy human ventricular model,[24] of which I_{Ks} models were replaced by the following equations based on WT-I_{Ks} or WT/G229D-I_{Ks} obtained by electrophysiologic recordings.

For both WT-I_{Ks} and WT/G229D-I_{Ks}:

$$I_{Ks} = G_{Ks} \cdot (1 + 0.6/(1 + (3.8 \cdot 10^{-5}/[Ca^{2+}]_i)^{1.4})) \cdot x_{s1} \cdot x_{s2} \cdot (V_m - E_{Ks})$$

$$dx_{s1}/dt = (x_{s,\infty} - x_{s1})/tau_{x,s1}$$

$$dx_{s2}/dt = (x_{s,\infty} - x_{s2})/tau_{x,s2}$$

For WT-I_{Ks}:

$$x_{s,\infty} = 1/(1 + \exp(-(V_m + 28.8)/15.45))$$

$$tau_{x,s1} = 326.9 + 0.4/(2.326 \cdot 10^{-4} \cdot \exp((V_m + 65.5)/17.8) + 1.292 \cdot 10^{-3} \cdot \exp(-(V_m + 227.2)/230))$$

$$tau_{x,s2} = 5/(0.01 \cdot \exp((V_m - 50)/100) + 0.0193 \cdot \exp(-(V_m + 66.54)/155))$$

For WT/G229D-I_{Ks}:

$$x_{s,\infty} = 0.85/(1 + \exp(-(V_m + 82.8)/41.72))$$

$$tau_{x,s1} = 326.9 + 0.4/(2.326 \cdot 10^{-4} \cdot \exp((V_m + 119.5)/17.8) + 1.292 \cdot 10^{-3} \cdot \exp(-(V_m + 281.2)/230))$$

$$tau_{x,s2} = 5/(0.01 \cdot \exp((V_m - 50)/100) + 0.0193 \cdot \exp(-(V_m + 66.54)/155))$$

where G_{Ks} (mS/μF) is the maximum conductance of I_{Ks}; $[Ca^{2+}]_i$ (mM) is the intracellular Ca^{2+} concentration; x_{s1} and x_{s2} are the activation and deactivation gates, respectively, for I_{Ks}; V_m (mV) is the transmembrane potential; E_{Ks} (mV) is the reversal potential for I_{Ks}; $x_{s,\infty}$ is the steady-state value of both x_{s1} and x_{s2} gates; and $tau_{x,s1}$ and $tau_{x,s1}$ are the time constant of x_{s1} and x_{s2} gates, respectively. The values of G_{Ks} for atrial and ventricular models were 0.0136 and 0.0034 mS/μF, respectively, because the WT-I_{Ks} with these values faithfully reproduced the same current amplitudes as the original I_{Ks} models.[23,24]

To obtain the ventricular transmural gradient, we defined endocardial, midmyocardial, and epicardial layers of thickness (0.6, 6.0, and 2.4 mm, respectively[25]) and then incorporated transmural differences in ion channels and intracellular Ca^{2+} dynamics according to the original code.[24] Pacing stimuli of 2-ms duration and strength twice diastolic threshold were applied transmembranously to the endocardial end at a cycle length of 1000 ms. To obtain ECGs similar to those recorded from left precordial leads, a unipolar recording electrode was located 2 cm above the epicardial end of the tissue. Transmural conductivity in the extracellular space was set to 2.36 mS/cm and that in the intracellular space for endocardial and midmyocardial layers and for epicardial layer were set to 0.38 and 0.29 mS/cm, respectively.[26] Other model parameters and the numerical approach have been described elsewhere.[25]

Statistical analysis

All data are given as mean ± SEM. Differences between two groups were examined by independent Student *t* test. $P < .05$ was considered significant.

Results

Genetic analysis

We identified three heterozygous mutations in 3 of 30 probands with juvenile-onset AF (10%): *SCN5A*-M1875T, *KCNJ2*-M301K, and *KCNQ1*-G229D. Details of patients with *SCN5A*-M1875T and *KCNJ2*-M301K have been reported previously.[5,6] The third *KCNQ1* mutation, a single-base substitution at nucleotide 686 (c.686G > A),

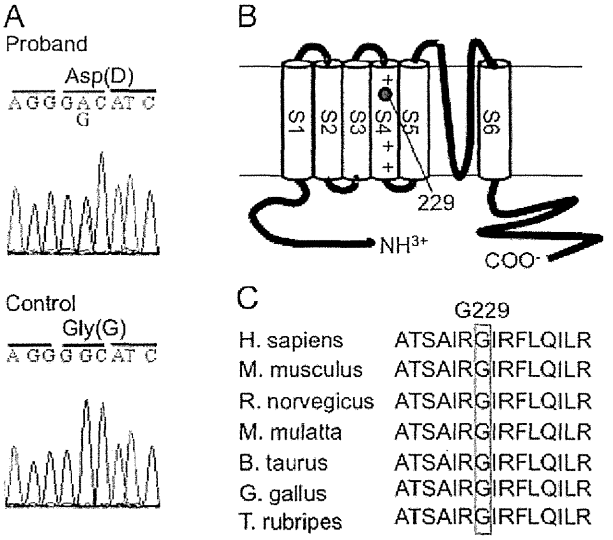


Figure 1 Genetic analysis of the proband. **A:** Electropherograms of *KCNQ1* gene showing a mutation, p.G229D (c.686G > A) in the proband. **B:** Topology of Kv7.1 encoded by KCNQ1. G229 is located in the fourth transmembrane segment (S4), known as the voltage sensor. **C:** Alignment of Kv7.1 sequence showing conservation of glycine at position 229 (G229) across species.

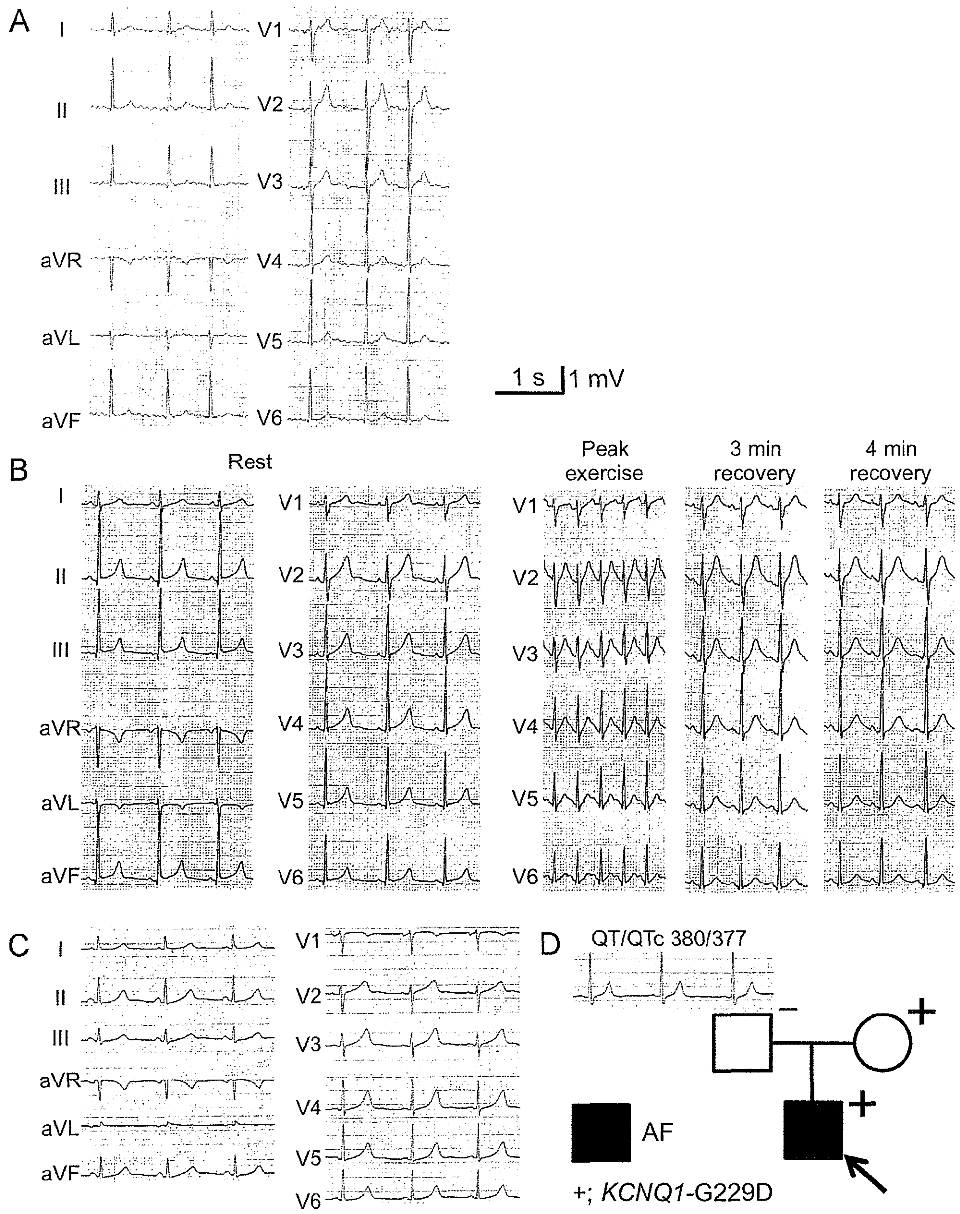


Figure 2 Clinical characteristics. Twelve-lead ECGs of the proband, indicating atrial fibrillation (AF) at the first detection (**A**) and sinus rhythm at rest and during exercise tolerance test after catheter ablation (**B**). **C:** ECG of the proband's mother. **D:** Pedigree and ECG (V_5 lead) of the proband's father. Males and female are represented as *squares* and *circle*, respectively. *Arrow* indicates the proband. +/− symbols indicate the presence/absence of the *KCNQ1*-G229D variant. *Filled symbols* indicate the development of AF.

causes an amino acid change from glycine to aspartic acid at position 229 in the Kv7.1 potassium channel (Figure 1A). Gly-229 resides in the fourth transmembrane segment (S4), which is known as a voltage sensor (Figure 1B). Alignment of the Kv7.1 amino acid sequence (Figure 1C) demonstrated that the glycine at position 229 is conserved in several species, suggesting its importance at this position. G229D was absent in 400 Japanese control alleles and has not been reported according to the NHLBI Exome Sequencing Project (ESP), Exome Variant Server (http://evs.gs.washington.edu/EVS/). In the proband, we did not find any mutations in other candidate genes described in the Methods.

Clinical characteristics

KCNQ1-G229D mutation was identified in a 16-year-old boy who was diagnosed as AF at the age of 16 (QT/QTc [corrected QT interval] 380/429 ms; Figure 2A). Cardiovascular and blood examination including thyroid hormone were all normal. He took propranolol hydrochloride 30 mg/day and digoxin 0.125 mg/day for rate control and bepridil 100 mg/day for pharmacologic cardioversion, but they failed to maintain his sinus rhythm. Moreover, cardioversion did not restore sinus rhythm.

As the next step of treatment, he underwent radiofrequency catheter ablation therapy (pulmonary vein isolation). After the therapy, his AF did not recur without any antiarrhythmic agents for 20 months. Eighteen months after therapy, exercise tolerance test was performed (Figure 2B). The QTc interval both at rest and at 4 minutes of recovery after exercise showed borderline criteria of QT prolongation (QT/QTc from 415/452 ms to 372/480 ms).

The G229D mutation was also identified in the boy's asymptomatic mother. Her ECG at rest showed borderline criteria of QT prolongation (QT/QTc 460/468 ms; Figure 2C). There was no family history of AF (Figure 2D).

Functional analysis

To elucidate the genetic effect of G229D mutant, we conducted functional characterization by using a heterologous expression system. As shown in Figure 3A, coexpression of *KCNQ1*-WT (0.5 μg) with *KCNE1* (0.5 μg) produced a slowly activating outward WT-I_{Ks} on depolarization to 30 mV from the holding potential of –80 mV. In contrast, transfection of *KCNQ1*-G229D (0.5 μg), coexpressed with equimolar *KCNE1*, produced an instantaneously activated G229D-I_{Ks} that did not deactivate after repolarization to –50 mV (Figure 3B). Moreover, the cells coexpressing WT and mutant (0.25 μg each) channels with *KCNE1* displayed both instantaneous (indicated by an arrow) and time-dependent activated WT/G229D-I_{Ks} with deactivation process (Figure 3C), suggesting that coexpression of WT results in an intermediate functional phenotype. HMR1556 (an I_{Ks} blocker; 1 μM) completely inhibited all conducted I_{Ks} currents (indicated by red traces in Figures 3A through 3C).

Activating currents were then divided into instantaneous and time-dependent components by measuring the instantaneous current level at 10 ms after the depolarization pulse and the time-dependent current as a difference between current level at 10 ms after depolarization and at the end of depolarizing duration (steady state (indicated by arrowhead in Figure 3C). As summarized in bar graphs of Figure 3D, instantaneous components of WT/G229D-I_{Ks} were 171.9 ± 28.3 pA/pF, which was significantly larger than those in WT-I_{Ks} (40.8 ± 14.1 pA/pF, $P < .001$). In contrast, time-

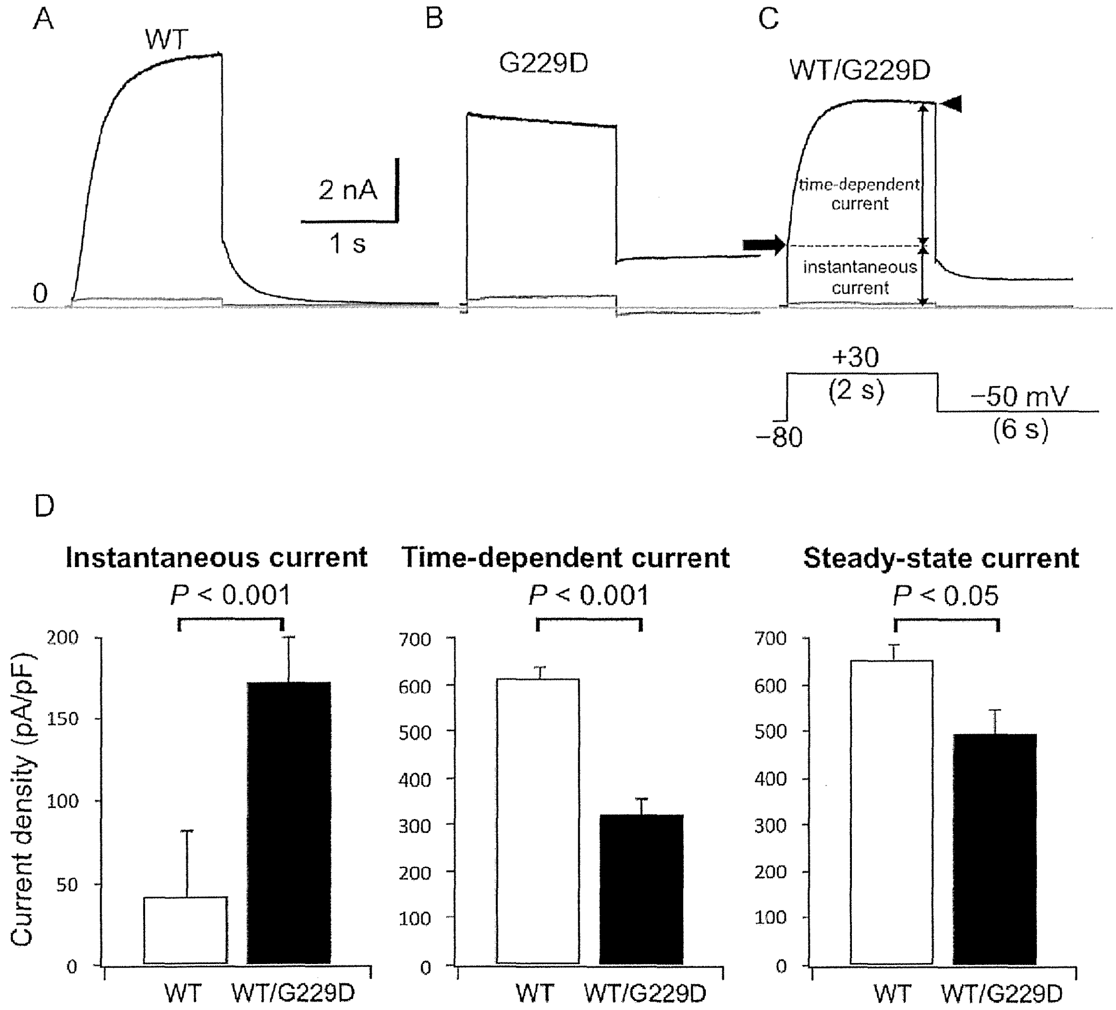


Figure 3 G229D mutation drastically alters properties of reconstituted I_{Ks} current. Whole-cell Kv7.1 currents recorded from CHO cells expressing wild type (WT; **A**), G229D (**B**), and WT/G229D (**C**). Current was elicited by 2-second voltage step from a holding potential of –80 mV to 30 mV before *(black trace)* and after *(red trace)* application of HMR 1556 (1 μM). *Blue lines* indicate the zero current level. *Inset (lower right)* shows the voltage application protocol. **D:** Bar graphs of current densities summarized from multiple experiments for instantaneous *(left)*, time-dependent *(middle)*, and steady-state *(right)* currents during voltage step to 30 mV. Instantaneous current level was measured 10 ms after depolarization *(arrow)*. Time-dependent current was estimated as a difference between an instantaneous current level *(arrow)* and that at the end of depolarizing pulse (steady-state: *arrowhead*). *White bars* indicate data from WT (n = 19). *Black bars* indicate data from WT/G229D (n = 19).

dependent and steady-state currents were significantly larger in WT-I_{Ks} than WT/G229D-I_{Ks} (Figures 3C and 3D).

Figures 4A depicts two families of I_{Ks} current traces elicited by 2-second depolarizing voltage-clamp steps from a holding potential of –80 mV to various test potentials and following repolarization to –50 mV (inset shows voltage-step protocol): left, WT-I_{Ks} and right, WT/G229D-I_{Ks}. Again, WT/G229D-I_{Ks} displayed an instantaneous activation (indicated by arrow in Figure 4A), which was followed by time-dependent slow activation (intermediate phenotype). On repolarization to –50 mV, tail currents did not completely deactivate. As shown in the lower panels, both WT and WT/G229D currents were entirely inhibited by HMR1556 (1 μM).

Figures 4B shows peak tail current–voltage relationships for WT-I_{Ks} and WT/G229D-I_{Ks} obtained from multiple experiments. Tail current densities were larger at test potentials between –130 and –40 mV in WT/G229D (filled circles) than in WT (open circles). At more depolarizing test pulses (20 to 50 mV), however, they were reversely smaller in WT/G229D than in WT.

In Figure 4C, voltage-dependent activations of WT and WT/G229D tail currents were evaluated by fitting to a Boltzmann equation. In WT/G229D, the voltage dependence for I_{Ks} activation was significantly shifted to more hyperpolarized potentials (from –15.1 ± 1.4 mV to –50.8 ± 7.8 mV).

Because deactivation of WT/G229D-I_{Ks} was extremely slow at –50 mV (Figures 3C and 4A), in the following experiments (Figure 5A) we measured deactivation kinetics (tau_{fast} and tau_{slow}) at –120 mV after 2-second depolarization to 30 mV (left trace, WT; right trace, WT/G229D). As summarized in the bar graphs shown in Figure 5B, compared to WT, rates for deactivation in WT/G229D were significantly slower (tau_{fast} 77.3 ± 3.6 ms vs 115.0 ± 8.9 ms; tau_{slow} 270.9 ± 124.9 ms vs 1716.3 ± 110.9 ms).

In the next series of experiments, we used an atrial AP waveform to elicit current activation. WT/G229D-I_{Ks} thus conducted large outward currents (Figure 6A, red trace). As summarized in the bar graphs shown in Figure 6B, integral current densities were significantly larger in WT/G229D than WT, suggesting a gain-of-function effect by G229D mutation.

Because the instantaneous current component of WT/G229D-I_{Ks} could influence the resting membrane potential, we measured the resting potential of CHO cells expressing various constructs (Figure 6C). Resting membrane potentials were –4.6 ± 1.9 mV in nontransfected cells and –41.3 ± 2.2 mV in cells expressing WT channels. In contrast, those

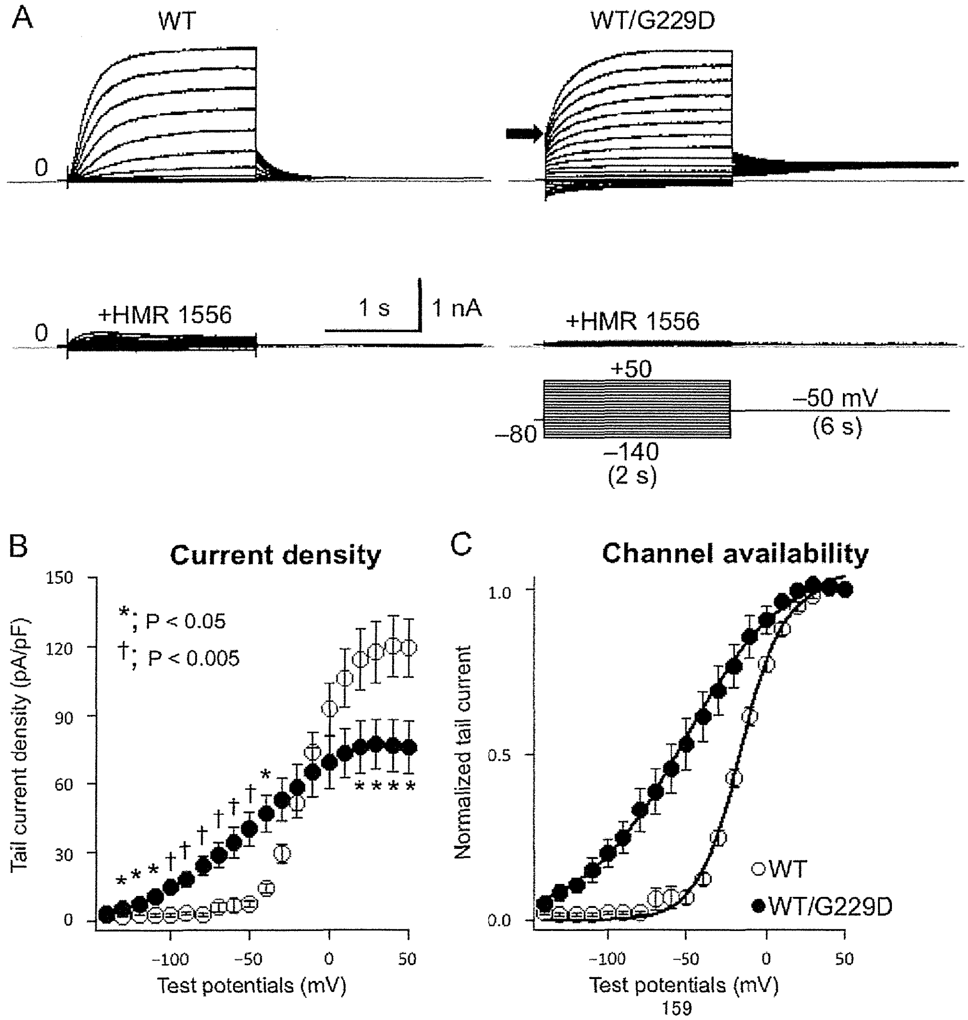


Figure 4 Electrophysiologic properties of wild type (WT) and WT/G229D. **A:** Two representative sets of current traces recorded from CHO cells expressing WT *(left)* and WT/G229D *(right)*. Currents were elicited by 2-second depolarizing voltage-clamp steps from a holding potential of –80 mV to various test potentials (from –140 to 50 mV) before *(top)* and after *(bottom)* application of HMR1556 (1 μM). *Arrow* in the *right panel* indicates the instantaneous current level at 50-mV test potential. *Blue lines* indicate the zero current level. *Inset (lower right)* shows the voltage application protocol. **B:** Tail current–voltage relationships for WT (*open circles*, n = 7) and WT/G229D (*filled circles*, n = 6). Tail current densities are plotted as a function of test potentials. *Vertical lines* through symbols indicate the standard error. **C:** Current–voltage relationships for normalized tail currents in WT (*open circles*, n = 7) and WT/G229D (*filled circles*, n = 6). *Vertical lines* through symbols indicate the standard error.

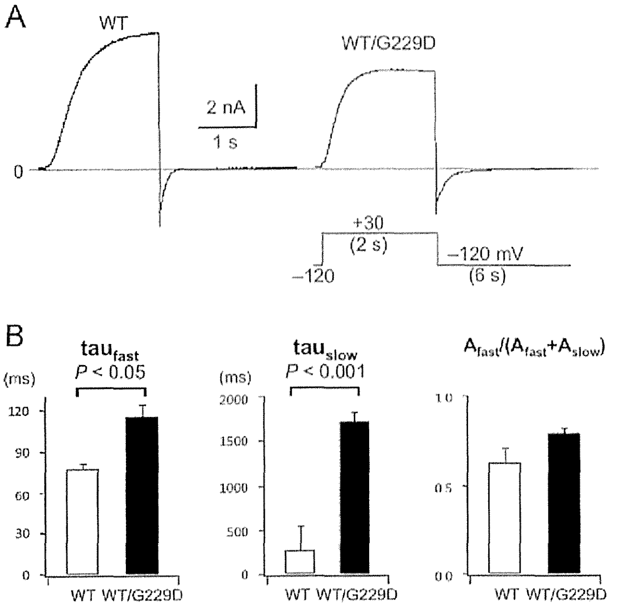


Figure 5 Coexpression with G229D slows down the deactivation process. **A:** Current traces from CHO cells expressing wild type (WT; *left*) and WT/G229D (*right*). Currents were elicited by a 2-second depolarizing voltage step to 30 mV followed by repolarization to −120 mV to obtain a completely deactivated tail current. *Blue line* indicates the zero current level. *Inset (lower right)* shows the voltage application protocol. **B:** Bar graphs showing averaged tau_{fast} *(left)*, tau_{slow} *(middle)*, and $A_{fast}/(A_{fast} + A_{slow})$ *(right)*. *White bars* indicate data from WT (n = 8). *Black bars* indicate date from WT/G229D (n = 10).

expressing WT/G229D or G229D channels showed significantly more negative resting potentials (–74.6 ± 1.9 mV and –74.3 ± 6.2 mV), which were closer to the calculated equilibrium potential of potassium ion (–89.9 mV in the present experimental condition). Therefore, the negative shift of resting potentials may be due to constitutive opening of WT/G229D or G229D channels.

Computer simulation

To explore the cellular mechanisms by which the G229D mutation manifested AP shortening in atrial but not ventricular myocytes, we performed a computer simulation study by using both atrial and ventricular myocytes and 1D myocardial model (Figure 7). Based on the I_{Ks} obtained by electrophysiologic recording (Figures 4 and 5), we numerically reproduced both WT and WT/G229D current traces (Figure 7A), the current–voltage relationship curves (Figure 7B), and the normalized activation curves (Figure 7C). The numerically reproduced I_{Ks} were incorporated into the human atrial and ventricular myocyte models (Figures 7D and E, respectively). WT/G229D-I_{Ks} was markedly larger than WT-I_{Ks} in both atrial and ventricular cell models. Because of the difference in the contribution of I_{Ks} to AP formation, the mutation markedly shortened AP duration in the atrial but not the ventricular myocyte model. Indeed, the numerically reproduced I_{Ks} did not shorten the QT interval in the 1D model under 1-Hz pacing (Figure 7F), consistent with the ECG phenotype of the proband.

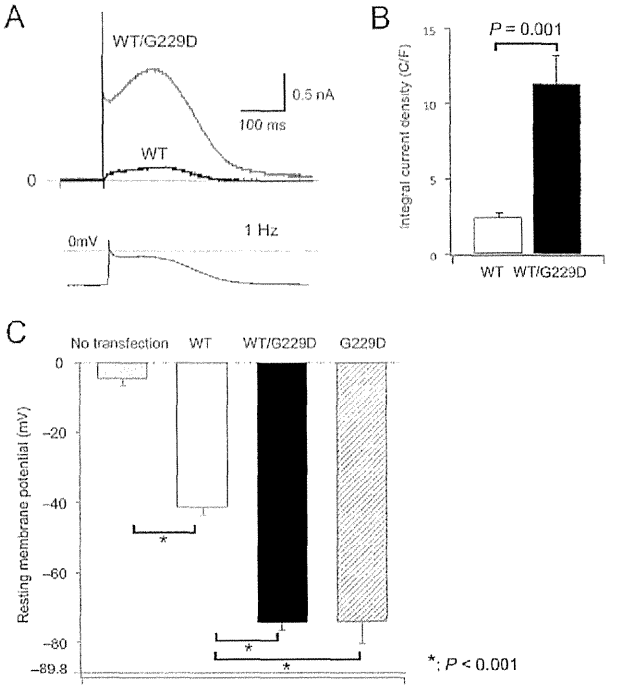


Figure 6 **A, B:** Atrial action potential clamp recording. **A:** Whole-cell clamp was conducted by using an atrial action potential waveform (indicated in *inset* below the traces). Two representative current traces recorded from cells expressing WT *(black trace)* and WT/G229D *(red trace)*. *Blue line* indicates the zero current level. **B:** Bar graph showing averaged integral current densities measured by the area enclosed by the current curves. *White bar* indicates data from WT (n = 10). *Black bar* indicates data from WT/G229D (n = 11). **C:** Bar graph showing averaged resting membrane potentials from CHO cells under four different conditions: no transfection of constructs (*dot bar,* n = 23), transfected with WT (*white bar,* n = 25), WT/G229D (*black bar,* n = 15), and G229D (*striped bar,* n = 12). *Blue line* indicates the equilibrium potential of potassium ion in the present experimental condition.

Discussion

In the present study, we described a novel missense *KCNQ1* mutation, G229D, in a juvenile-onset AF patient. The proband's AF, which started at the age of 16 years, was refractory to bepridil. Radiofrequency catheter ablation therapy was effective in maintaining his sinus rhythm. In sinus rhythm, he showed borderline QT prolongation. His mother carried the same heterozygous mutation and also showed borderline QT prolongation.

G229D mutant I_{Ks} reconstituted in CHO cells displayed unique functional properties: a time-independent component (instantaneous current) and slow deactivation. In detail, (1) instantaneous component of WT/G229D was significantly larger than that of WT; (2) the tail current density of WT/G229D was larger at test potentials between –130 and –40 mV; (3) WT/G229D produced a negative shift in the voltage dependence of half-maximal activation (–35.2 mV); (4) the deactivation of WT/G229D was significantly slower than that of WT; (5) a large integral current density of WT/G229D was indeed induced by the atrial AP clamp experiment; and (6) computational AP simulations suggest WT/G229 selectively shortens the atrial AP. Taken together, these results are

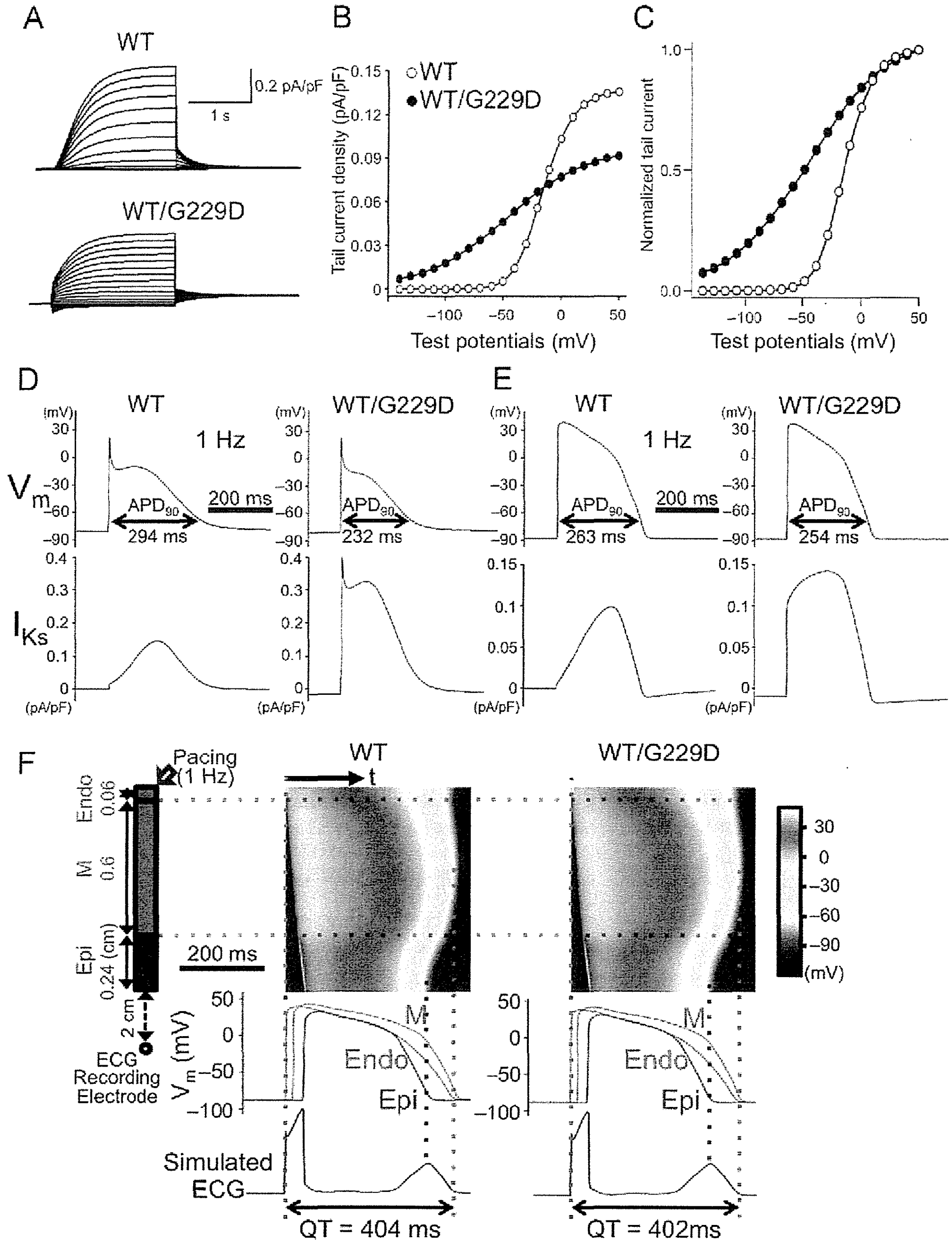


Figure 7 Computer simulation study using both atrial and ventricular myocytes and one-dimensional (1D) myocardial model. Numerically reproduced current traces (**A**), tail current–voltage relationships (**B**) and normalized activation curves (**C**) of wild-type (WT) model *(open circles)* and WT/G229D model *(filled circles)*. **D:** Atrial action potential (AP) and I_{Ks} in the computer models of human atrial myocyte with WT-I_{Ks} or WT/G229D-I_{Ks}. **E:** Ventricular AP and I_{Ks} in the computer models of human ventricular myocyte with WT-I_{Ks} or WT/G229D-I_{Ks}. **F:** Transmural heterogeneity of AP and the simulated ECG in the 1D ventricular model.

consistent with a view that the mutation caused gain-of-function effects on I_{Ks}, thereby shortening atrial refractoriness and increasing susceptibility to AF.

Regarding *KCNQ1* mutations associated with juvenile-onset AF, five mutations (S140G, V141M, S209P, R231C, R231H) were previously identified.[4,16,19–21] Chen et al[4] first reported the *KCNQ1*-S140G mutation that potentiated I_{Ks}, especially the component of instantaneous activation. Later in 2005, Hong et al[16] reported the *KCNQ1*-V141M mutation in a baby with AF and an abnormally short QT interval. They also described a large instantaneous activation of V141M-I_{Ks}. Das et al[19] then reported a heterozygous *KCNQ1*-S209L mutation. This mutation also showed an instantaneous opening when expressed as WT/S209L-I_{Ks}, a significantly negative shift of half-maximal activation voltage (–42.4 mV), and slow current deactivation.

More recently, Bartos et al[20,21] reported *KCNQ1*-R231C and R231H mutations in families with AF and mild QT prolongation. These mutations showed marked instantaneous activation and significantly negative shift in half-maximal activation (–30 to –40 mV). In addition, recent extensive mutagenesis experiments and the structural model of KCNQ1 protein[15] suggest that residues S140, E160, R237, and R231 closely associate with one another in the closed state. Substitution of amino acid at any of the charged S140, E160, R237, or R231 residues was shown to disrupt KCNQ1 deactivation ("lock" the I_{Ks} channel in the open state).[15,20] This suggests that these residues are critical for normal KCNQ1 channel closing and that G229D might also disrupt the interaction among these residues.

Our biophysical assessment revealed that the function of G229D resembled that of R231C and R231H mutations,[20,21] which reside near glycine at 229. However, the G229D mutation was somewhat different in that it caused a borderline QT prolongation. Indeed, it appeared not to affect the ventricular AP while markedly shortening the atrial AP. Our computer simulation 1D model (Figure 7F) partially explained these apparently different effects of G229D for the first time. Regarding the pharmacologic treatment of AF in the proband, a low concentration of pure I_{Ks} blocker would be a potential for restoring sinus rhythm without considerable prolongation of QT interval. In fact, Courtney et al[27] recently reported the enhanced sensitivity of *KCNQ1* gain-of-function mutation (S140G) for HMR1556, a pure I_{Ks} blocker, compared to that of WT channels.

Study limitations

In the present study, we used a heterologous expression system to assess functional modulation by the *KCNQ1* mutation. However, the environment of this system is different from that of cardiac myocytes or whole heart. Therefore, our data might not always explain electrophysiologic modulation in the whole heart.

Conclusion

We identified a novel *KCNQ1*-G229D mutation in a patient with juvenile-onset AF. In the heterozygous condition, the mutation changed I_{Ks} channel kinetics and showed a gain-of-function modulation of I_{Ks}. In the computer simulation model, it markedly shortened the atrial AP duration, suggesting the tendency toward AF.

Acknowledgments

We are grateful to the Japanese juvenile-onset AF families for their willingness to participate in this study. We thank Dr. Daniel C. Bartos (Department of Physiology, University of Kentucky, Lexington, USA) for the AP waveform and Arisa Ikeda, Kazu Toyooka, and Aya Umehara for excellent technical assistance.

References

1. Kannel WB, Benjamin EJ. Status of the epidemiology of atrial fibrillation. Med Clin North Am 2008;92:17–40, ix.
2. Brand FN, Abbott RD, Kannel WB, Wolf PA. Characteristics and prognosis of lone atrial fibrillation. 30-year follow-up in the Framingham Study. JAMA 1985;254:3449–3453.
3. Sinner MF, Ellinor PT, Meitinger T, Benjamin EJ, Kaab S. Genome-wide association studies of atrial fibrillation: past, present, and future. Cardiovasc Res 2011;89:701–709.
4. Chen YH, Xu SJ, Bendahhou S, et al. KCNQ1 gain-of-function mutation in familial atrial fibrillation. Science 2003;299:251–254.
5. Makiyama T, Akao M, Shizuta S, et al. A novel SCN5A gain-of-function mutation M1875T associated with familial atrial fibrillation. J Am Coll Cardiol 2008;52:1326–1334.
6. Hattori T, Makiyama T, Akao M, et al. A novel gain-of-function KCNJ2 mutation associated with short-QT syndrome impairs inward rectification of Kir2.1 currents. Cardiovasc Res 2012;93:666–673.
7. Ravn LS, Aizawa Y, Pollevick GD, et al. Gain of function in I_{Ks} secondary to a mutation in KCNE5 associated with atrial fibrillation. Heart Rhythm 2008;5: 427–435.
8. Lundby A, Ravn LS, Svendsen JH, Hauns S, Olesen SP, Schmitt N. KCNE3 mutation V17M identified in a patient with lone atrial fibrillation. Cell Physiol Biochem 2008;21:47–54.
9. Yang Y, Xia M, Jin Q, et al. Identification of a KCNE2 gain-of-function mutation in patients with familial atrial fibrillation. Am J Hum Genet 2004;75:899–905.
10. Ellinor PT, Nam EG, Shea MA, Milan DJ, Ruskin JN, MacRae CA. Cardiac sodium channel mutation in atrial fibrillation. Heart Rhythm 2008;5:99–105.
11. Olson TM, Alekseev AE, Liu XK, et al. Kv1.5 channelopathy due to KCNA5 loss-of-function mutation causes human atrial fibrillation. Hum Mol Genet 2006;15:2185–2191.
12. Gollob MH, Jones DL, Krahn AD, et al. Somatic mutations in the connexin 40 gene (GJA5) in atrial fibrillation. N Engl J Med 2006;354:2677–2688.
13. Hodgson-Zingman DM, Karst ML, Zingman LV, et al. Atrial natriuretic peptide frameshift mutation in familial atrial fibrillation. N Engl J Med 2008;359: 158–165.
14. Mahida S, Lubitz SA, Rienstra M, Milan DJ, Ellinor PT. Monogenic atrial fibrillation as pathophysiological paradigms. Cardiovasc Res 2011;89:692–700.
15. Restier L, Cheng L, Sanguinetti MC. Mechanisms by which atrial fibrillation-associated mutations in the S1 domain of KCNQ1 slow deactivation of I_{Ks} channels. J Physiol 2008;586:4179–4191.
16. Hong K, Piper DR, Diaz-Valdecantos A, et al. De novo KCNQ1 mutation responsible for atrial fibrillation and short QT syndrome in utero. Cardiovasc Res 2005;68:433–440.
17. Lundby A, Ravn LS, Svendsen JH, Olesen SP, Schmitt N. KCNQ1 mutation Q147R is associated with atrial fibrillation and prolonged QT interval. Heart Rhythm 2007;4:1532–1541.
18. Otway R, Vandenberg JI, Guo G, et al. Stretch-sensitive KCNQ1 mutation A link between genetic and environmental factors in the pathogenesis of atrial fibrillation? J Am Coll Cardiol 2007;49:578–586.
19. Das S, Makino S, Melman YF, et al. Mutation in the S3 segment of KCNQ1 results in familial lone atrial fibrillation. Heart Rhythm 2009;6:1146–1153.
20. Bartos DC, Duchatelet S, Burgess DE, et al. R231C mutation in KCNQ1 causes long QT syndrome type 1 and familial atrial fibrillation. Heart Rhythm 2011;8:48–55.
21. Bartos DC, Anderson JB, Bastiaenen R, et al. A KCNQ1 mutation causes a high penetrance for familial atrial fibrillation. J Cardiovasc Electrophysiol 2013;24:562–569.
22. Bosch RF, Nattel S. Cellular electrophysiology of atrial fibrillation. Cardiovasc Res 2002;54:259–269.
23. Courtemanche M, Ramirez RJ, Nattel S. Ionic mechanisms underlying human atrial action potential properties: insights from a mathematical model. Am J Physiol 1998;275:H301–H321.
24. O'Hara T, Virag L, Varro A, Rudy Y. Simulation of the undiseased human cardiac ventricular action potential: model formulation and experimental validation. PLoS Comput Biol 2011;7:e1002061.
25. Tsuji-Wakisaka K, Akao M, Ishii TM, et al. Identification and functional characterization of KCNQ1 mutations around the exon 7-intron 7 junction affecting the splicing process. Biochim Biophys Acta 2011;1812:1452–1459.
26. Yamada KA, Kanter EM, Green KG, Saffitz JE. Transmural distribution of connexins in rodent hearts. J Cardiovasc Electrophysiol 2004;15:710–715.
27. Campbell CM, Campbell JD, Thompson CH, Vanoye CG, George AL Jr. Selective targeting of gain-of-function KCNQ1 mutations predisposing to atrial fibrillation. Circ Arrhythm Electrophysiol 2013;6:960–966.