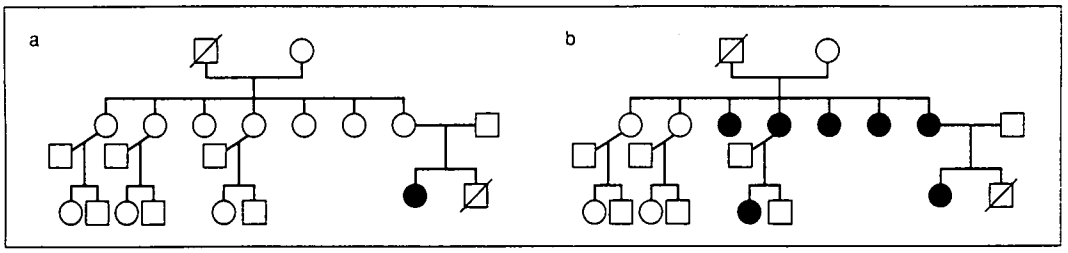

図1　不完全浸透(incomplete penetrance)
心電図診断(a)では発端者(●)のみが異常であるが，遺伝子解析(b)を行うと家系内にさらに6人の変異キャリアがいることがわかる．〔文献5)より引用〕

不完全浸透の詳細な分子メカニズムは完全には解明されていないが，LQTSの臨床像に大きな影響を与える要因の一つとして，遺伝子変異のもたらす機能異常の重症度が挙げられる．機能異常が強い変異のキャリアではほとんどが顕性化すると思われるが，逆に機能異常が軽微であれば，なんらかの代償機構によって不顕性化することが推測される．この代償機構は「再分極予備能(repolarization reserve)」[6]と呼ばれる概念である．心筋の再分極相は遅延整流K電流(IKr・IKs)一過性外向きK電流(transient outward K current：Ito)をはじめとする多くのK電流の働きによって成り立っているため，もし低K血症や薬剤などによってIKrが多少抑制されても，その他のK電流などによる代償機構によって再分極が保たれ，活動電位持続時間が正常範囲に収まるというものである．この再分極予備能になんらかの個人差があれば，薬剤感受性の個人差や不完全浸透の原因となる可能性がある．LQTS遺伝子や薬物代謝遺伝子以外の未知の遺伝的素因(modifier gene)がQT時間や不整脈発症に影響を与える可能性も否定できない．

後天性LQTSのもう一つの機序として挙げられるのは，薬剤のターゲット分子(すなわちイオンチャネル)の遺伝的な個人差である．一般に，「遺伝子変異」は母集団のゲノムの1%以下にしか認められない遺伝子バリエーションで，多くは機能異常を伴い，遺伝病の原因となりうるものである．一方，1%以上の頻度でみられる遺伝子バリエーションは「遺伝子多型」と呼ばれる．遺伝子多型のなかには，遺伝子のアミノ酸配列や発現量を変化させ機能変化を伴うもの，軽微な影響しかないもの，逆にそれらに全く無関係のものなど，その表現型は様々である．遺伝子多型のうち，一塩基遺伝子多型(single nucleotide polymorphism：SNP)は，様々な病態発症や薬剤感受性を規定する修飾因子の候補として注目されている．LQTS関連遺伝子にも遺伝子多型の存在が報告されており，そのうちのいくつかは通常は臨床的には明らかにならないが，チャネルレベルでは軽微(subclinical)な機能異常を有するものが最近判明してきた[7,8]．Yangらは薬剤誘発性TdP症例92例と，同様の治療を受けてもQT延長を示さなかったコントロールについて，LQT関連遺伝子の検索を行い，いくつかの変異とSNPを見いだした[9]．五つのLQTS関連遺伝子のアミノ酸置換型SNPのうち，後天性LQTSに報告されたものと人種差や薬剤感受性の変化を認めたものの部位を記載する(**図2**)．

Brugada症候群

Brugada症候群は，心電図右側胸部誘導のCoved型・Saddle-back型ST上昇を特徴とする，器質的心疾患がない心室細動(特発性心室細動)である．中年男性の突然死の原因として知られる「ぽっくり病」と同一の疾患と考えられている．これまで*SCN5A*が唯一の原因遺伝子として同定されてきたが[10]，*SCN5A*変異が同定されるのは患者全体の20～30%に過ぎず，他の原因遺伝子の存在が推測されていた．最近，Weissらはglycerol-3-phosphate dehydrogenase like(GPD1-L)遺伝子にNaチャネルのトラフィッキング(小胞体で翻訳・修飾を受けた蛋白質が，細胞膜まで到達する過程)を阻害する変異を同定した[11]．また，QT短縮を合併したBrugada症候群家系に

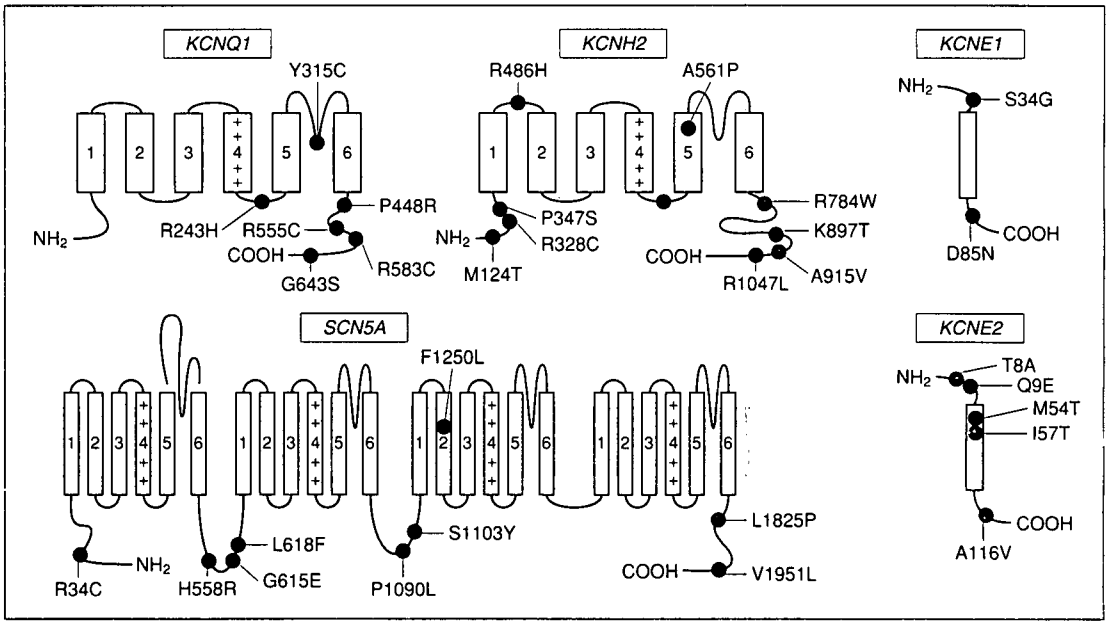

図2　薬剤誘発性QT延長症候群に認められたLQT関連遺伝子の変異とアミノ酸置換型SNP
五つのLQTS遺伝子の膜貫通領域をボックスで示し，SNPの位置を●で示す．SNPの多くはチャネル機能に重大な影響を及ぼす可能性の高い膜貫通領域以外に比較的多い傾向がある．

Caチャネルβ2サブユニット(*CACNB2b*)の変異も報告された[12]．QT延長症候群の遺伝子分類と同様に，*SCN5A*，*GPD1-L*，*CACNB2b*責任遺伝子とするBrugada症候群をそれぞれBrS1，BrS2，BrS3と分類することが提案されている．一方，全く無症状でありながらBrugada型の心電図を呈する無症候性Brugada症候群の病因には未解明の部分が多い．

QT短縮症候群(short QT syndrome；SQTS)

近年，QT間隔の短縮(QTc＜360 msec)を特徴とする遺伝性不整脈SQTSに三つの原因遺伝子(*KCNH2*・*KCNQ1*・*KCNJ2*)が同定された[13]．これらはそれぞれLQT2・LQT1・LQT7の原因遺伝子と同一であるが，SQTS変異はKチャネル電流を増加する(gain-of-function)点でLQTSと異なる．

カテコラミン誘発性多形性心室頻拍(catecholaminergic polymorphic VT；CPVT)

CPVTは，運動や精神的ストレスによって二方向性心室頻拍や多形性心室頻拍が誘発される先天性不整脈で，小児や青年期の突然死の原因の一つである．約9%の症例に家族歴を認め，常染色体優性遺伝形式をとることが多い．原因遺伝子は，心筋のCaハンドリングを制御する筋小胞体Ca遊離チャネル(リアノジン受容体*RyR2*)とカルセクエストリン2(*CASQ2*)である[14,15]．運動や精神的ストレスにより上昇したカテコラミンが，本来細胞内Ca濃度が低い拡張期においても筋小胞体から異常なCaリークをきたし，細胞内Ca過負荷による遅延後脱分極を生じ致死的不整脈が誘発される．

催不整脈性右室異形成(arrhythmogenic right ventricular cardiomyopathy；ARVC)

ARVCは右室の線維脂肪化による拡大・機能低下と，心室性不整脈を特徴とする心筋症である．5,000人に1人の割合で発症し，若年の突然死の原因の約20%を占める[16]．また運動中に急変することが多く，運動選手の突然死の原因としても注目されている．原因遺伝子として，*RyR2*・*TGFβ3*・デスモプラキン・プラコフィリン2などが同定されている[17]．

家族性心臓ブロック

家族性心臓ブロックは，PCCD(progressive cardiac conduction defect)またはLev-Lenègre病とも呼ばれ，刺激伝導系の進行性障害により脚ブロック・完全房室ブロックをきたす．loss-of-functionの機能異常を有する*SCN5A*変異が報告されている[18]．同じloss-of-functionのNaチャネル変異が，なぜBrugada症候群とPCCDという異なる臨床像を示すのかに関してはまだ不明な点が多い．

家族性心房細動(familial atrial fibrillation)

心房細動(AF)は日常診療で見受けられる最も頻度の高い不整脈の一つだが，AFのうち明らかな原因のないlone AFが15～30%を占める[19]．また，AFの5%に家族歴があることからも[20]，遺伝的な要因の関与が示唆されていた．最近家族性AFの原因遺伝子として，Kチャネル*KCNQ1*・*KCNE2*・*KCNJ2*遺伝子にgain-of-functionを呈する変異が同定された[21]．これらの変異は活動電位持続時間を短縮するため，心房の有効不応期の短縮によってリエントリーが生じやすくなり，AFが出現すると考えられる．

先天性洞不全症候群(sick sinus syndrome；SSS)

SSSは通常，加齢や基礎心疾患に合併して発生することが多いが，稀に先天性の症例もみられる．Schultz-Bahrらは優性遺伝形式の先天性SSS家系にペースメーカチャネル*HCN4*の変異を同定した[22]．一方Bensonらは，劣性遺伝形式の先天性SSS家系に*SCN5A*変異を同定した[23]．洞結節細胞の脱分極を担う最も重要な電流はCa電流であるが，Na電流は全く関与していないのか，洞結節細胞内にNaチャネルは発現していないのかということについては，これまで論議が多かった．しかし，最近の報告の示すとおり，Naチャネル洞結節の中心部には発現されていないものの，周辺部分に確かに発現しており[24]，*SCN5A*が先天性SSSの原因遺伝子であることと必ずしも矛盾しない．われわれも最近，機能異常が比較的軽微な二つの異なる*SCN5A*変異を父方と母方から別々に引き継いで発症した，複合ヘテロによる先天性SSS症例の一家系を報告した(論文投稿中)．

おわりに

不整脈の遺伝子診断の歴史はまだ浅く，多くの候補遺伝子の中から，変異を同定するにはかなりの時間を要するのが現状であり，遺伝子解析が不整脈の日常診療に役立つようになるためには，いくつかの技術革新が必要である．しかし，遺伝子診断は，発端者や同胞の遺伝子異常を明らかにするだけでなく，将来的には，遺伝情報を基に薬剤投与前から不整脈発生の危険性を予知し薬剤性不整脈を回避する「テーラーメイド医療」の実現に道を開くと考えられる．今後，変異や多型を含めた多くの遺伝子解析データと機能解析データの集積によって，不整脈の原因を解明し有効な治療法を開発することが期待される．

文　献

1) Kass RS, Moss AJ：Long QT syndrome：novel insights into the mechanisms of cardiac

arrhythmias. J Clin Invest 112 : 810-815. 2003
2) Mohler PJ, Schott JJ, Gramolini AO, et al : Ankyrin-B mutation causes type 4 long-QT cardiac arrhythmia and sudden cardiac death. Nature 421 : 634-639. 2003
3) Plaster NM, Tawil R, Tristani-Firouzi M, et al : Mutations in Kir2.1 cause the developmental and episodic electrical phenotypes of Andersen's syndrome. Cell 105 : 511-519. 2001
4) Roden DM : Drug-induced prolongation of the QT interval. N Engl J Med 350 : 1013-1022. 2004
5) Priori SG, Napolitano C, Schwartz PJ : Low penetrance in the long-QT syndrome : clinical impact. Circulation 99 : 529-533. 1999
6) Roden DM : Taking the "idio" out of "idiosyncratic" : predicting torsades de pointes. Pacing Clin Electrophysiol 21 : 1029-1034. 1998
7) Donger C, Denjoy I, Berthet M, et al : KVLQT1 C-terminal missense mutation causes a forme fruste long-QT syndrome. Circulation 96 : 2778-2781. 1997
8) Napolitano C, Schwartz PJ, Brown AM, et al : Evidence for a cardiac ion channel mutation underlying drug-induced QT prolongation and life-threatening arrhythmias. J Cardiovasc Electrophysiol 11 : 691-696. 2000
9) Yang P, Kanki H, Drolet B, et al : Allelic variants in long-QT disease genes in patients with drug-associated Torsades de Pointes. Circulation 105 : 1943-1948. 2002
10) Chen Q, Kirsch GE, Zhang D, et al : Genetic basis and molecular mechanism for idiopathic ventricular fibrillation. Nature 392 : 293-296. 1998
11) Weiss R, Barmada MM, Nguyen T, et al : Clinical and molecular heterogeneity in the Brugada syndrome : A novel gene locus on chromosome 3. Circulation 105 : 707-713. 2002
12) Antzelevitch C, Pollevick GD, Cordeiro JM, et al : Loss-of-function mutations in the cardiac calcium channel underlie a new clinical entity characterized by ST-segment elevation, short QT intervals, and sudden cardiac death. Circulation 115 : 442-449. 2007
13) Brugada R, Hong K, Dumaine R, et al : Sudden death associated with short-QT syndrome linked to mutations in HERG. Circulation 109 : 30-35. 2004
14) Priori SG, Napolitano C, Memmi M, et al : Clinical and molecular characterization of patients with catecholaminergic polymorphic ventricular tachycardia. Circulation 106 : 69-74. 2002
15) Lahat H, Pras E, Olender T, et al : A missense mutation in a highly conserved region of CASQ2 is associated with autosomal recessive catecholamine-induced polymorphic ventricular tachycardia in Bedouin families from Israel. Am J Hum Genet 69 : 1378-1384. 2001
16) Gemayel C, Pelliccia A, Thompson PD : Arrhythmogenic right ventricular cardiomyopathy. J Am Coll Cardiol 38 : 1773-1781. 2001
17) Bauce B, Basso C, Rampazzo A, et al : Clinical profile of four families with arrhythmogenic right ventricular cardiomyopathy caused by dominant desmoplakin mutations. Eur Heart J 26 : 1666-1675. 2005
18) Schott JJ, Alshinawi C, Kyndt F, et al : Cardiac conduction defects associate with mutations in SCN5A. Nat Genet 23 : 20-21. 1999
19) Wiesfeld AC, Hemels ME, Van Tintelen JP, et al : Genetic aspects of atrial fibrillation. Cardiovasc Res 67 : 414-418. 2005
20) Darbar D, Herron KJ, Ballew JD, et al : Familial atrial fibrillation is a genetically heterogeneous disorder. J Am Coll Cardiol 41 : 2185-2192. 2003
21) Chen Y-H, Xu S-J, Bendahhou S, et al : KCNQ1 gain-of-function mutation in familial atrial fibrillation. Science 299 : 251-254. 2003
22) Schulze-Bahr E, Neu A, Friederich P, et al : Pacemaker channel dysfunction in a patient with sinus node disease. J Clin Invest 111 : 1537-1545. 2003
23) Benson DW, Wang DW, Dyment M, et al : Congenital sick sinus syndrome caused by recessive mutations in the cardiac sodium channel gene (SCN5A). J Clin Invest 112 : 1019-1028. 2003
24) Tellez JO, Dobrzynski H, Greener ID, et al : Differential expression of ion channel transcripts in atrial muscle and sinoatrial node in rabbit. Circ Res 99 : 1384-1393. 2006

致死性不整脈の遺伝子基盤に関するイオンチャネルのゲノム薬理学的研究
—イオンチャネルの遺伝子多型と致死性不整脈—

Ion channel gene polymorphisms underlying lethal arrhythmias

蒔田 直昌[*1] 佐々木 孝治[*2] 清水 渉[*3] 堀江 稔[*4]

Key words：不整脈，イオンチャネル，遺伝子変異，遺伝子多型，*SCN5A*

はじめに

心筋虚血・心筋症などの心疾患や一部の薬剤は，致死性不整脈発生の危険因子であり，患者の予後を大きく左右する。しかしこれらの危険因子を持つ人が全て致死性不整脈を発生するわけではなく，不整脈の発生しやすさは何らかの遺伝的素因によって規定されている可能性がある。その遺伝的素因の候補と考えられるのは，薬物代謝酵素やイオンチャネルの遺伝子多型であり，最近特に一塩基遺伝子多型(SNP)が注目されている。イオンチャネルの遺伝子変異や多型のもたらす機能異常が軽微であれば，通常は心電図異常が顕性化することはないが，そこに薬物などのあらたな外的刺激が加わると，二次的にチャネルの機能異常が顕性化し,致死性不整脈が発生する可能性がある。

本研究の目的は，イオンチャネルの遺伝子多型や変異と薬剤感受性や不整脈の重症度との因果関係を明らかにすることである。ある特定の遺伝子多型や不顕性変異を有する群が持たないものに比べ致死性不整脈の発生率が高い，あるいは薬物感受性が高いということが判明すれば，その遺伝子多型・変異を持つ患者に対してはあらかじめ危険性の高い薬剤の投与を回避することが可能になる。さらに本研究では，わが国を含めた東アジアで特に罹患率が高い，心臓突然死の一因であるBrugada症候群の遺伝子基盤を明らかにする。症候性・無症候性Brugada症候群患者の臨床像と遺伝型をSNP解析によって相関解析を行い，Brugada症候群の重症度を規定する遺伝的素因を明らかにする。

方　法

1. 正常人・患者のゲノム解析

血縁関係のない正常ボランティア120人，先天性QT延長症候群（LQTS）45例，後天性LQTS 4例，無症候性Brugada症候群 30例，有症候性Brugada 37例，進行性心臓ブロック（PCCD）2例，心房停止3例，洞不全症候群（SSS）4例の末梢血よりPuregene genomeキット（Gentra）を用いてゲノムDNAを抽出した。先天性不整脈の原因としてすでに明らかにされている5つの心筋イオンチャネル遺伝子，*SCN5A*，*KCNQ1*，*KCNH2*，*KCNE1*，*KCNE2*のエクソンとその周囲をPCRで増幅し，アクリルアミドゲルによるsingle-strand conformational polymorphism（SSCP）法，ABI prism 310 genetic analyzer（Applied Biosystems）を用いた直接シークエンス法で，遺伝子変異・多

*1 Makita Naomasa　北海道大学大学院医学研究科　循環病態内科学
*2 Sasaki Koji　同　上
*3 Shimizu Wataru　国立循環器病センター　心臓血管内科
*4 Horie Minoru　滋賀医科大学　呼吸循環器内科

型をスクリーニングした。一部の症例ではギャップジャンクションであるコネキシン40（Cx40）・コネキシン43（Cx43）および，ペースメーカチャネルHCN4の遺伝子解析も行った。変異・SNP解析は直接シークエンスまたはTaqMan法を用いて行った。

2. 遺伝子変異の機能解析

同定した*SCN5A*変異については，Quick Change Mutagenesisキット（Stratagene）を用いて，ヒトNaチャネルcDNA（hNav1.5）を鋳型に変異cDNAを作成し，pcDNA3.1発現ベクター（Invitrogen）にサブクローニング後，哺乳類細胞tsA201細胞に一過性トランスフェクションした。24時間後，パッチクランプ法で全細胞Na電流を測定し，既報の方法で変異Naチャネルの機能を解析した[1]。フレカイナイドは灌流液中に2〜10 μM投与し，平衡状態に達してから測定した。

結　果

1. イオンチャネル遺伝子解析（表1，図1）

イオンチャネルの遺伝子変異スクリーニングに

表1　本研究で同定したイオンチャネル遺伝子変異

遺伝子	変異	疾患	機能異常	発表論文
KCNQ1	A178T	LQT1		既報
KCNQ1	G269S	LQT1		既報
KCNQ1	G325R	LQT1		既報
KCNQ1	R591H	LQT1		既報
KCNH2	E58N	LQT2		既報
KCNH2	insIAQ in 82-84	LQT2		既報
SCN5A	E1784K	LQT3	GOF	既報
SCN5A	L1825P	後天性 LQTS	LOF+GOF	1)
SCN5A	A1746T	Brugada 症候群+LQT3		
SCN5A	Q55X	Brugada 症候群	無機能	1)印刷中
SCN5A	P1725R	Brugada 症候群		
SCN5A	D356N	Brugada 症候群	無機能	2)
SCN5A	R367H	Brugada 症候群+心房停止	無機能	3)
SCN5A	K1527R+A1569P	無症候性 Brugada 症候群	LOF	4)
SCN5A	L212P	心房停止	LOF	5)
SCN5A	R219H	洞不全症候群	無機能	論文作成中
SCN5A	M1880V+801-803S	先天性洞不全症候群	LOF	論文作成中
SCN5A	c.5290delG	進行性心筋伝導障害	無機能	

LOS: loss-of-function, GOF: gain-of-function

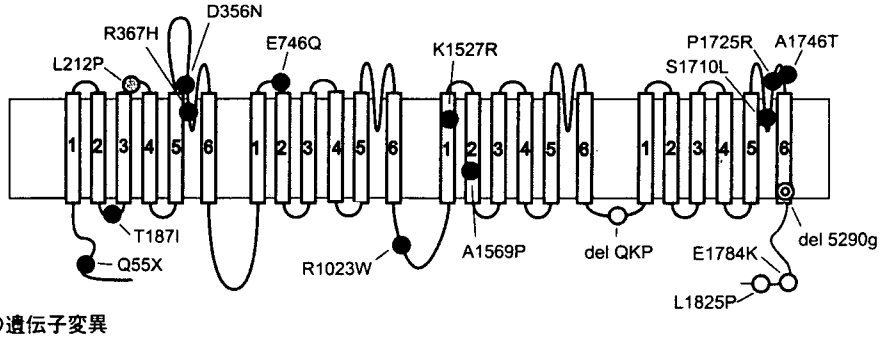

図1　*SCN5A*の遺伝子変異

本研究で同定した新規*SCN5A*変異。○はLQT3．●はBrugada症候群．◎はPCCD．●は心房停止に認められた*SCN5A*変異の部位を示す。Q55X＋E746Qは同一alleleに存在するdouble mutationであった。同様にK1527R＋A1569Pは無症候性Brugada症候群症例の同一allele上に存在するdouble mutationだった。c. del5290GはPCCDとBrugada症候群の合併例に認められた1塩基欠損変異である。

よって先天性LQTS 4例に*KCNQ1*変異を，2家系に*KCNH2*変異を同定した。また，先天性LQTS 2例（うち1例Brugada症候群合併例），後天性LQTS 1例，無症候性Brugada症候群 1例，有症候性Brugada症候群 5例（うち1例LQTS合併例，1例心房停止合併例），心房停止 2例，SSS 2例，PCCD1例に*SCN5A*の遺伝子変異を同定した。

*SCN5A*変異のうち，Q55X[11]，D356N[2]，R367H[3]，R219H，c.5290delGは無機能チャネルであった。先天性・後天性LQTS症例は典型的な3型LQTS（LQT3）様の遅延電流を示した。また，無症候性Brugada症候群症例に見られた変異K1527R＋A1569Pは同一エクソン上の二重変異であり，Na電流を低下させる特徴的なチャネル機能異常（loss-of-function）を示した[4]。また，先天性SSSに認められたM1880V＋801-803Sは複合ヘテロ変異（compound heterozygosity）であった。L1825Pは典型的なLQT3様のgain-of-functionとloss-of-functionを併せ持っていた[1]。

L212Pは11歳男児の心房停止症例に認められた*SCN5A*変異である[5]。心房停止は徐脈，P波の消失，房室接合部補充収縮を特徴とするまれな不整脈である。患児は4歳のときに心房細動とSSSの診断を受けているが，失神の既往はない。父方祖父の兄弟が突然死している。入院時心電図はP波がなく，心室性補充収縮を伴う房室結節調律で，

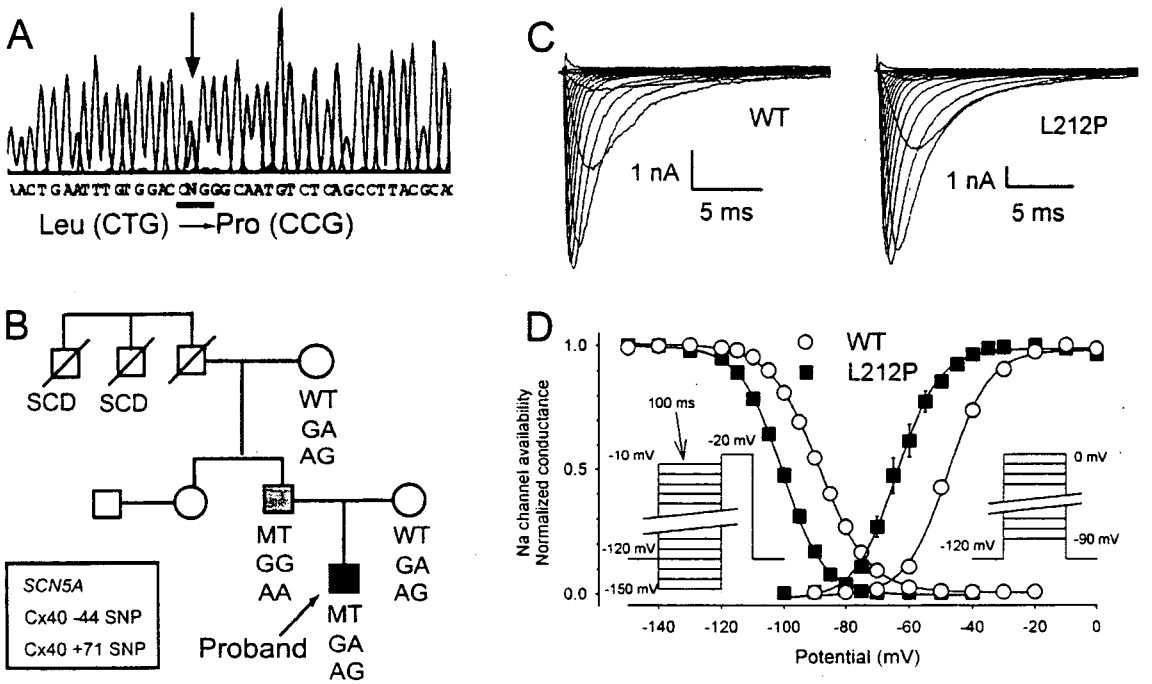

図2　心房停止に認められた*SCN5A*変異とCx40のSNP

A. L212P変異のゲノムシークエンス

B. 心房停止家系。遺伝子型は上から*SCN5A*，Cx40 −44，Cx40 +71を示す。WT＝正常，MT＝L212P，SCD＝心臓突然死。父親はL212Pのキャリアだが，心電図は正常で，Cx40は通常型。

C. tsA201細胞に発現させた正常およびL212PNa^+チャネルのwhole cell電流。

D. 不活性化（左），活性化（右）の膜電位依存性曲線。L212Pチャネルは不活性化・活性化がともに大きく過分極側にシフトしている。

QRS幅・QT時間は正常だった。遺伝子解析では，ペースメーカチャネルHCN4とCx40, Cx43に変異はなかったが，*SCN5A*にミスセンス変異L212Pを認めた（**図2A**）。父親は無症状で心電図もほぼ正常であったがL212Pのキャリアであった。母親の*SCN5A*は正常であった（**図2B**）。パッチクランプ法で測定したL212PのNa電流は，活性化・不活性化に大きな異常を示した(**図2C, 2D**)。このような重篤な機能異常を示す変異L212Pを有したために患児は心房停止を示した一方で，父親はこの変異を有しながら心電図はほぼ正常であった。Cx40の転写調節領域に存在するSNPが心房停止や心房細動の誘発性に関与しているとの報告[6]から，正常日本人コントロールと本家系においてCx40のSNPを解析した。正常日本人99人でCx40の遺伝子多型をTaqMan法で解析すると，その頻度は−44GG/＋71AAが63%，−44GA/＋71AGが32%，−44AA/＋71GGが5%であった[5]。心筋細胞を使ったプロモータ解析から，−44AA/＋71GGは−44GG/＋71AAに比較してCx40の転写活性が有意に低いことが明らかになっている。本家系では，患児と母親はヘテロのCx40多型をもつのに対して，父親は通常型遺伝型であった。したがって，患児は父親から*SCN5A*変異L212Pを，母親からCx40多型を受け継ぎ，心房停止に至ったと推測される。

2. *SCN5A*の遺伝子多型（SNP）

以前から報告されている*SCN5A*のSNPのほかに，2種の新規SNPとしてD1114N・L1988Rを同定した。これらのSNPについて正常人，有症候性・無症候性Brugada症候群患者のallele頻度をTaqMan PCR法および直接sequence法で検討した。L1988Rは日本人特異的な*SCN5A*のSNPとして本研究ではじめて同定した（**表2**）。そのアレル頻度を正常人，症候性・無症候性Brugada症候群，QT延長症候群で比較検討した。**表2**のとおり，L1988Rのアレル頻度は，正常人2.5%，Brugada症候群5.1%，無症候性Brugada症候群0%，QT延長症候群0.67%であり，各群間の統計学的検定（χ^2検定）では有意差は認められなかった。同様に，D1114NやH558RなどのSNP頻度にも疾患群間に有意差は認められなかった。

表2　日本人特異的新規SNP L1988Rハプロタイプ
正常(T/T)・L1988Rホモ(G/G)およびヘテロ(G/T)のアレル頻度を示す。

	T/T	G/T	G/G
正常	114	8	0
症候性 BS	62	6	0
無症候性 BS	60	0	0
QT 延長症候群	150	2	0

次に，**図3A**に示した9種類のミスセンスSNPのうち，R53Cを除く8種類のSNPについて，その機能を正常と比較検討した。それぞれのSNP Naチャネルの機能をパッチクランプ法で解析した。**図3B**は，電気生理学特性のうち，不活性化・活性化の膜電位依存性を示す。4種類のSNPチャネルの膜電位依存性に関しては，正常チャネルと有意な差を認めなかった。不活性からの回復や，遅い不活性化など，その他のチャネル特性にも有意差はなかった。また，Brugada症候群の心電図異常を増強するⅠc群抗不整脈薬フレカイナイド（2〜10μM）を用い，ブロックの濃度依存性，トニックブロック・頻度依存性ブロックの程度などを比較したが，これに関してもSNPとWTチャネルの間に有意差は認められなかった。これらの事実は，*SCN5A*のミスセンスSNPは少なくとも単一では，生理学的・薬理学的な機能変化をもたらす遺伝的修飾因子とはいえないことを示唆している。

考　察

イオンチャネルの遺伝子多型に関する研究は世界的にもまだ緒についたばかりであるが，心筋Kチャネル遺伝子*KCNE2*のSNPが薬剤誘発性二次

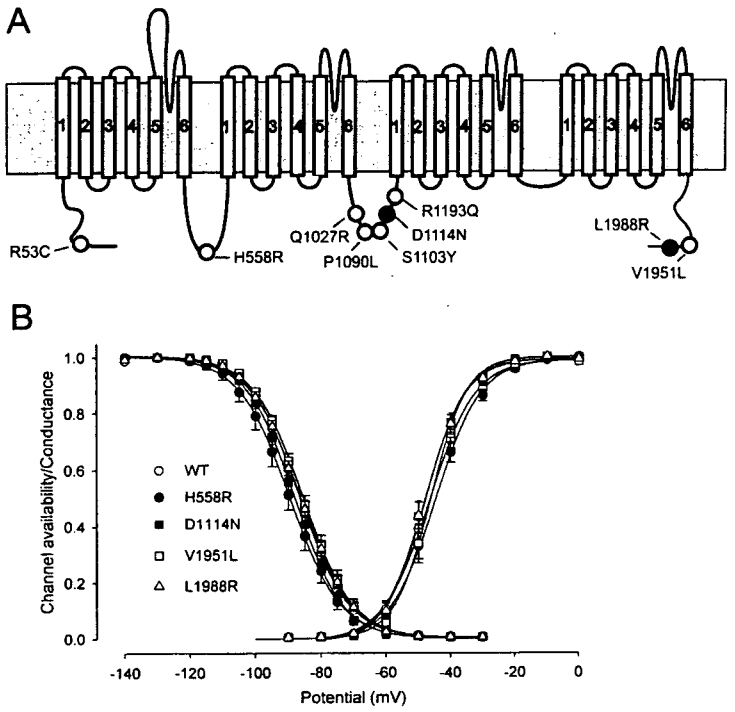

図3 *SCN5A*のSNPの部位と機能解析

A. アミノ酸配列を変えるSNPのみを列挙した。○は既報のSNP，●は本研究で同定した新規SNPである。図1の変異が*SCN5A*全体に分布し，特定の部位に集中していないのに対して，SNPはすべて細胞内ドメインで，特にドメイン2～3間のループに集中している。

B. 正常（WT）および4種類のミスセンスSNPチャネルを作成し，tsA201細胞に発現させwhole cell Na電流を測定した。不活性化（左）および活性化（右）の膜電位依存性を示す（mean±SE）。WTに比べて統計学的有意差は見られない（n=8～22）。

性LQTSの病態に深くかかわっていると報告されて以来，急速に注目を浴びている研究領域である[7]。我々は，イオンチャネルの不顕性変異や遺伝子多型が，先天性不整脈の病因であるだけではなく，基礎心疾患を有する患者の薬剤感受性や不整脈の重症度を規定する，重要な修飾因子（modifier）であるとの仮説を立て，イオンチャネル病の遺伝子解析・機能解析を行った。

本研究において18種類のイオンチャネル遺伝子変異を同定した。そのうち7個は既報の変異であったが，残りの11個は新規変異であった（**表1**）。このうち機能が全くない*SCN5A*変異を5種類見いだした。同じ無機能チャネルでありながら，その臨床像は全く異なっていた。このことから，致死性不整脈の原因となるイオンチャネル病の病像は，原因となるチャネル遺伝子変異だけで一義的に決まるのではなく，ホルモンや自律神経の影響，チャネルの機能に影響を与える他の修飾分子・環境因子の影響も受けていることが推測された。この修飾因子（modifier）のひとつの候補が遺伝子多型である。

本研究では4種類の*SCN5A*のミスセンスSNPについて電気生理学的機能異常・薬理学的異常の有無を検索したが，どれも正常*SCN5A*とほとんど違いは見られなかった（**図3B**）。Ackermanらも最

近，39種類に及ぶ*SCN5A*のSNPの人種差を報告したが，単独で機能異常を示すものはなかった[8]。最も頻度の多いSNPであるH558Rも単独で機能異常を示すことはないが，最近，他の変異チャネルM1766L・T512I・R282HやSNP Q1077delに多彩な機能修飾を与えるという報告が相次いでいる。また最近，*SCN5A*上流調節領域に，アジア人に特異的で転写活性を低下させるハプロタイプHapBの存在が報告された。このハプロタイプが直ちにBrugada症候群の病態や人種差の直接原因であると結論づけることはできないが，伝導障害を含めたさまざまな不整脈の分子基盤あるいは修飾因子のひとつとして今後の研究が注目される[9]。このように不整脈の病態の修飾因子としてSNPの役割が次々に明らかになっており，今後これらの遺伝子変異の機能評価をする際には，SNPの存在を考慮することが必要であると思われる。

また本研究では，心房停止症例で認めた*SCN5A*変異L212Pに対して，Cx40のSNPの遺伝型がmodifierとして作用していると推測しているが，この仮説はさらに多くの患者や家系で再評価する必要があるのはいうまでもない。しかし，Cx40のSNPは心房細動の誘発性に関与しているとの報告もすでにあり[10]，Cx40が心房内の刺激伝導系の修飾因子として作用し，そこに別の器質的な機能異常が加わることによって臨床的に心房の伝導障害が顕性化するとの考察は可能である。一方，無症候性Brugada症候群に関しては，*SCN5A*遺伝子異常との因果関係や，*SCN5A*異常の有無が予後判定の基準になりうるのかどうかは，症例数が少なく本研究では明らかにすることはできなかった。しかし無症候性Brugada症候群に*SCN5A*の二重変異[4]が同定されたのは世界的にも初めてであり，重要な知見である。

本研究では，不整脈の発生しやすさや薬剤感受性を規定する修飾遺伝子やSNPは特定するにはいたらなかった。しかしイオンチャネル病の遺伝子解析やSNP解析をさらに進めることによって，致死性不整脈の遺伝子基盤やその修飾遺伝子を明らかにするだけでなく，将来，「不整脈のオーダーメイド医療」を構築するための重要なエビデンスになることが期待される。

謝　辞

本研究の遂行にあたり，研究助成を頂きました臨床薬理研究振興財団に深く御礼申し上げます。

文　献

1) Makita N, Horie M, Nakamura T, et al. Drug-induced long-QT syndrome associated with a subclinical SCN5A mutation. *Circulation* 2002; **106**: 1269-1274.

2) Makiyama T, Akao M, Tsuji K, et al. High risk for bradyarrhythmic complications in patients with Brugada syndrome caused by *SCN5A* gene mutations. *J Am Coll Cardiol* 2005; **46**: 2100-2106.

3) Takehara N, Makita N, Kawabe J, et al. A cardiac sodium channel mutation identified in Brugada syndrome associated with atrial standstill. *J Intern Med* 2004; **255**: 137-142.

4) Yokoi H, Makita N, Sasaki K, et al. Double SCN5A mutation underlying asymptomatic Brugada syndrome. *Heart Rhythm* 2005; **2**: 285-292.

5) Makita N, Sasaki K, Groenewegen WA, et al. Congenital atrial standstill associated with coinheritance of a novel SCN5A mutation and connexin 40 polymorphisms. *Heart Rhythm* 2005; **2**: 1128-1134.

6) Groenewegen WA, Firouzi M, Bezzina CR, et al. A cardiac sodium channel mutation cosegregates with a rare connexin40 genotype in familial atrial standstill. *Circ Res* 2003; **92**: 14-22.

7) Sesti F, Abbott GW, Wei J, et al. A common polymorphism associated with antibiotic-induced cardiac arrhythmia. *Proc Natl Acad Sci U S A* 2000; **97**: 10613-10618.

8) Ackerman MJ, Splawski I, Makielski JC, et al. Spectrum and prevalence of cardiac sodium channel variants among black, white, Asian, and Hispanic individuals: implications for arrhythmogenic susceptibility and Brugada/long QT syndrome genetic testing. *Heart Rhythm* 2004; **1**: 600-607.

9) Bezzina CR, Shimizu W, Yang P, et al. Common sodium channel promoter haplotype in asian subjects underlies variability in cardiac conduction. *Circulation* 2006; **113**: 338-344.

10) Firouzi M, Ramanna H, Kok B, et al. Association of human connexin40 gene polymorphisms with atrial vulnerability as a risk factor for idiopathic atrial fibrillation. *Circ Res* 2004; **95**: e29-33.

11) Makita N,Sumitomo N, Watanabe I, Tsutsui H Novel *SCN5A* mutation(Q55X) associated with age-dependent expression of Brugada syndrome presenting as neurally mediated syncope. *Heart Rhythm* 2007; **4**: 516-517.

A Functional Polymorphism in *COL11A1*, Which Encodes the α1 Chain of Type XI Collagen, Is Associated with Susceptibility to Lumbar Disc Herniation

Futoshi Mio, Kazuhiro Chiba, Yuichiro Hirose, Yoshiharu Kawaguchi, Yasuo Mikami, Takeshi Oya, Masaki Mori, Michihiro Kamata, Morio Matsumoto, Kouichi Ozaki, Toshihiro Tanaka, Atsushi Takahashi, Toshikazu Kubo, Tomoatsu Kimura, Yoshiaki Toyama, and Shiro Ikegawa

Lumbar disc herniation (LDH), degeneration and herniation of the nucleus pulposus of the intervertebral disc (IVD) of the lumbar spine, is one of the most common musculoskeletal diseases. Its etiology and pathogenesis, however, remain unclear. Type XI collagen is important for cartilage collagen formation and for organization of the extracellular matrix. We identified an association between one of the type XI collagen genes, *COL11A1*, and LDH in Japanese populations. *COL11A1*, which encodes the α1 chain of type XI collagen, was highly expressed in IVD, but its expression was decreased in the IVD of patients with LDH. The expression level was inversely correlated with the severity of disc degeneration. A single-nucleotide polymorphism (c.4603C→T [*rs1676486*]) had the most significant association with LDH ($P = 3.3 \times 10^{-6}$), and the transcript containing the disease-associated allele was decreased because of its decreased stability. These observations indicate that type XI collagen is critical for IVD metabolism and that its decrease is related to LDH.

Lumbar disc herniation (LDH), degeneration and herniation of the nucleus pulposus of intervertebral disc (IVD) of the lumbar spine, is one of the most common musculoskeletal diseases.[1–3] Its etiology and pathogenesis, however, remain unclear. Genetic factors have been implicated in the etiology of lumbar disc degeneration.[4,5] Genetic abnormalities of the extracellular matrix (ECM) are implicated in disc degeneration and LDH. Phenotypes of transgenic mice and human mutations underscore the candidacy of ECM genes as susceptibility genes for LDH.[6,7] Several researchers have reported the association of ECM protein genes, including genes for type IX collagen[8,9] and aggrecan,[10] with lumbar disc disease (LDD). We reported elsewhere that cartilage intermediate layer protein and asporin—ECM proteins highly expressed in IVD, as well as articular cartilage—are implicated in LDD.[11,12]

Type XI collagen is a cartilage-specific ECM protein important for cartilage collagen fibril formation and for ECM organization.[13–16] Type XI collagen is composed of three α-chains, α1(XI), α2(XI), and α3(II), which are encoded by *COL11A1*, *COL11A2*, and *COL2A1*, respectively. The three chains fold into triple-helical heterotrimers to form procollagen, which is secreted into the ECM, where it participates in fibril formation with other cartilage-specific collagens, type II and IX collagens.[13] Type XI collagen regulates the diameter of cartilage collagen fibrils. Its N-terminal noncollagenous region limits the appositional lateral growth of the fibril by blocking further accretion of type II collagen.[14,15] Chondrodysplasia in mouse (*cho*) is an autosomal recessive disorder due to a frame-shift mutation of *COL11A1*.[16] The collagen fibrils of *cho* mice are much thicker than normal.[16,17] Thus, type XI collagen has a critical role in the organization of the supramolecular architecture of cartilage collagen.

Type XI collagen is present in IVD, both in the annulus fibrosus and nucleus pulposus,[18] but its significance in LDH is not known. Type XI collagen is a quantitatively minor component of cartilage collagen fibrils, but it is essential for the interaction between proteoglycan (PG) aggregates and collagens. It binds with high affinity to PG, which is important in vivo for anchoring cartilage PG to the collagen fibrillar network.[19] Mutations in type XI collagen cause various types of chondrodysplasias in human, including Stickler syndrome type II (MIM #604841), Marshall syndrome (MIM #154780), and oto-spondylo-mega-epiphyseal dysplasia (MIM #215150). These disorders are collectively termed "type XI collagenopathies,"[20] and all are complicated by abnormalities of the spine, including narrowing of the IVD. In particular, patients with Stickler syndrome have spondylar abnormalities and Schmorl's node (disc herniation into the vertebral body).[21] These human mutations are in vivo evidence that type XI collagen is critical for IVD integrity; thus, the type XI collagen genes are good candidates for the gene that causes LDH.

From the Laboratories for Bone and Joint Diseases (F.M.; Y.H.; S.I.), Cardiovascular Diseases (K.O.; T.T.), and Statistics (A.T.), SNP Research Center, RIKEN, and Departments of Orthopaedic Surgery (F.M.; K.C.; Y.H.; M.K.; Y.T.) and Musculoskeletal Reconstruction and Regeneration Surgery (M. Matsumoto), School of Medicine, Keio University, Tokyo; Departments of Orthopaedic Surgery (Y.K.; T. Kimura) and Pathology (T.O.), Faculty of Medicine, University of Toyama, Toyama, Japan; Department of Orthopaedics, Graduate School of Medical Science, Kyoto Prefectural University of Medicine, Kyoto, Japan (Y.M.; M. Mori; T. Kubo); and Department of Orthopaedic Surgery, Keiyu Hospital, Yokohama, Japan (M.K.)

Received June 26, 2007; accepted for publication August 9, 2007; electronically published October 16, 2007.

Address for correspondence and reprints: Dr. Shiro Ikegawa, Laboratory for Bone and Joint Diseases, SNP Research Center, RIKEN, 4-6-1 Shirokanedai, Minato-Ku, Tokyo 108-8639, Japan. E-mail: sikegawa@ims.u-tokyo.ac.jp

Am. J. Hum. Genet. 2007;81:1271–1277. 0002-9297/2007/8106-0013$15.00
DOI: 10.1086/522377

Table 1. Clinical Characteristics of Subjects

Screening and Group	No. of Subjects	Age (years) Mean ± SD	Range	Male (%)	BMI[a]
1st:					
Case:					
LDD[b]	188	26.5 ± 10.4	13–74	40.0	21.0
LDH only	130	25.5 ± 6.9	13–66	54.0	21.1
Control	179	58.7 ± 11.7	23–81	6.0	23.0
2nd[c]:					
Case	359	41.4 ± 14.6	15–77	62.4	23.1
Control	286	69.6 ± 9.2	38–87	58.4	24.3
3rd[c]:					
Case	334	41.8 ± 15.1	11–83	61.3	23.4
Control	376	53.9 ± 9.7	13–86	47.6	22.2

[a] BMI calculated as body weight in kilograms divided by the square of height in meters.
[b] Includes disc degeneration only and LDH.
[c] Case group in the 2nd and 3rd screenings has LDH only.

Here, we present evidence that *COL11A1*, one of the type XI collagen genes, contributes to the genetic risk of LDH in Japanese. We have observed significant association between LDH and a functional SNP in *COL11A1* in independent Japanese populations. *COL11A1* was highly expressed in IVD, but its expression was decreased in the IVD of patients with LDH. *COL11A1* expression level was inversely correlated with the severity of disc degeneration in patients with LDH, and the transcript containing the disease-associated allele of the SNP was decreased.

Material and Methods

Study Population

All subjects were Japanese who were living in the middle part of the Honshyu island in Japan (table 1). They visited the participating hospitals and received medical examinations. For the initial screening, we recruited 188 case patients with LDD and 179 control subjects. The mean ages of the case and control groups were 26.5 and 58.7 years, respectively. The case group included 58 patients who had no herniation (disc degeneration only) and 130 patients with LDH. The mean age of the LDH case patients was 25.5 years. For the second screening (replication study), we recruited 359 patients with LDH and 286 control subjects. The mean ages of the case and control groups were 41.4 and 69.6 years, respectively. For the third screening, we recruited 334 patients with LDH and 376 control subjects. The mean ages of the case and control groups were 41.8 and 53.9 years, respectively. Subjects for the initial, second, and third screenings were recruited at the participating hospitals in the Toyama, Tokyo, and Kyoto areas, respectively. All LDH case patients had unilateral pain radiating from the back along the femoral or sciatic nerve to the corresponding dermatome of the nerve root with duration of >3 mo. Radiographic examination, including functional four-direction images and magnetic resonance imaging (MRI) (sagittal and axial images obtained with a 1.5-T imaging system), revealed positive findings indicating disc herniation. The degree of disc degeneration was evaluated by MRI and was scored according to Schneiderman's classification.[22] Of the affected individuals, 787 case patients underwent surgical treatment, and the other individuals with LDH were treated conservatively. All were followed up for >1 year. We excluded from the study individuals with spinal canal stenosis, spondylolisthesis, spondylosis, synovial cysts, spinal tumor, and trauma. We also excluded those who had occupational and/or habitual risk factors, such as heavy manual laborers, occupational drivers, and heavy smokers. We obtained informed consent from each subject, as approved by the ethical committees at the SNP Research Center of RIKEN and the participating hospitals.

Genotyping

We selected sequence variations of the type XI collagen genes (*COL11A1*, *COL11A2*, and *COL2A1*) for the first screening from the International HapMap Project database and JSNP Database. The SNPs covered >90% of the alleles with an r^2 value ≥0.8. We identified additional sequence variations in *COL11A1* by direct sequencing of a 230-kb region of DNA from 24 case patients. We extracted genomic DNA for genotyping from peripheral blood leukocytes of the subjects and genotyped SNPs as described elsewhere.[11,12]

Haplotype Structure and Statistical Analyses

We estimated haplotype frequencies, using the expectation-maximization algorithm and pairwise linkage-disequilibrium index (D' and Δ in 465 control individuals, as described elsewhere).[23] χ^2 tests were used to compare cases with controls for allelic and genotypic frequencies; the odds ratio (OR) and its 95% CI were calculated. We used a permutation test to adjust significance in the analysis of association between the *COL11A1* SNPs and LDH.[24] We performed 10^7 permutations of the cases and the controls. Bonferroni correction was applied when significance was adjusted for the number of SNPs genotyped. MRI data, real-time PCR data, and mRNA stability data were tested using Student's t test.

Analysis of COL11A1 *Expression*

We extracted and purified total RNAs and synthesized randomly primed cDNAs, using Multiscribe reverse transcriptase (PE Ap-

Table 2. Association between LDH and c.4603C→T (*rs1676486*) in *COL11A1*

Screening and Case Group	No. of Cases with Genotype CC	CT	TT	Total No. of Cases	No. of Controls with Genotype CC	CT	TT	Total No. of Controls	T Allele Frequency Case	Control	*P*	OR (95% CI)
1st:												
LDD[a]	85	86	17	188	99	67	13	179	.31	.26	.076	1.34 (.97–1.84)
LDH only	55	60	15	130	99	67	13	179	.34	.26	.020	1.51 (1.07–2.14)
2nd:												
LDH only	149	163	47	359	154	108	21	283	.35	.26	.00038	1.55 (1.21–1.97)

[a] Includes disc degeneration only and LDH.

Table 3. Polymorphisms in *COL11A1* and Their Association with LDH

Location in *COL11A1* and Nucleotide Sequence Change	Amino Acid Change	dbSNP	No. in the Three Genotype Groups[a]		Allelic Frequency		*P*[b]		OR (95% CI)[c]
			Case	Control	Case	Control	Allele	Genotype[d]	
IVS1:									
9284T→C	...	*rs1415359*	423/63/1	422/42/1	.07	.05	.068	.16	.69 (.47-1.03)
IVS6:									
82274A→C	...	...	437/49/1	424/38/1	.05	.04	.35	.61	.82 (.53-1.25)
IVS10:									
90221G→A	...	*rs945748*	426/62/1	414/48/1	.07	.05	.29	.54	.82 (.56-1.19)
IVS11:									
90406A→G	...	*rs3767272*	396/76/3	401/55/0	.09	.06	.032	.049	1.47 (1.03-2.10)
IVS20:									
104122A→T	...	*rs2622877*	438/47/2	400/46/0	.05	.05	.94	.38	.98 (.65-1.48)
IVS26:									
111262T→C	...	*rs2786125*	428/49/1	429/33/1	.05	.04	.11	.24	.70 (.45-1.08)
IVS41:									
146354T→C	...	*rs1012282*	425/62/1	415/47/1	.07	.05	.24	.47	1.26 (.86-1.84)
IVS42:									
165864A→C	...	*rs1841838*	381/104/3	374/84/6	.11	.1	.52	.27	1.10 (.82-1.47)
IVS44:									
169351A→G	...	*rs2126643*	378/100/3	373/79/6	.11	.1	.44	.23	1.12 (.84-1.51)
172702C→G	...	*rs3767273*	382/103/3	372/84/4	.11	.1	.41	.5	.88 (.66-1.19)
IVS50:									
192606G→A	...	*rs4908273*	231/211/43	271/167/23	.31	.23	.00023	.001	1.47 (1.20-1.80)
Exon 52:									
193817(c.3968)T→C	L1323P	*rs3753841*	193/230/65	238/187/38	.37	.28	.000081	.00041	1.47 (1.21-1.79)
IVS52:									
194187T→C	...	...	218/214/48	258/178/26	.32	.25	.00038	.0016	.69 (.57-0.85)
IVS54:									
200918A→G	...	*rs3767274*	399/73/4	367/86/5	.09	.1	.15	.34	.79 (.58-1.08)
206255G→T	...	*rs3767275*	457/30/0	442/15/1	.03	.02	.088	.068	.60 (.33-1.09)
208970T→A	...	*rs1676500*	443/45/1	425/33/1	.05	.04	.29	.53	1.27 (.81-1.99)
IVS58:									
218282C→G	...	...	431/46/1	430/32/1	.05	.04	.15	.32	.72 (.46-1.13)
Exon 62:									
219597(c.4603)C→T	P1535S	*rs1676486*	204/223/62	252/177/33	.35	.26	.000015	.000099	1.54 (1.27-1.88)
Exon 63:									
221284(c.4770)C→T	I1590I	*rs2229783*	169/236/83	214/201/47	.41	.32	.000028	.00017	1.49 (1.24-1.80)
IVS63:									
221659G→A	...	*rs1463048*	169/235/83	212/199/50	.41	.32	.000081	.00047	1.46 (1.21-1.76)
IVS65:									
225275T→A	...	*rs3753844*	207/223/55	239/186/33	.34	.28	.0014	.0056	1.38 (1.13-1.68)
Exon 67 (3′ UTR):									
230265C→T	...	*rs1031820*	443/45/1	430/33/1	.05	.04	.27	.5	.78 (.50-1.21)
230461A→G	...	...	439/45/1	429/33/0	.05	.04	.17	.3	.73 (.46-1.15)

NOTE.—The cDNA (accession number NM001854.2) and genomic DNA (accession numbers AC093150.4, AL627203.7, and AC099567.2) sequences of *COL11A1* are based on data from GenBank. The A of the ATG translation initiation codon in the reference sequence corresponds to position +1.

[a] Homozygote of the major allele/heterozygote/homozygote of the minor allele.

[b] By the χ^2 test.

[c] Calculated for the alleles.

[d] Calculated for the homozygotes of the major allele versus the heterozygotes and the homozygotes of the minor allele.

plied Biosystems). We performed quantitative real-time PCR using the ABI PRISM 7700 (Applied Biosystems) and QuantiTect SYBR Green PCR (QIAGEN) according to the manufacturer's instructions.

RNA Stability Assay

We amplified by PCR ~1,700-bp of *COL11A1* cDNA that contained the entire ORF. We cloned the *COL11A1* cDNA containing the associated SNP c.4603C→T into pCR-Blunt II-TOPO (Invitrogen) and confirmed the sequence of the inserts. Vectors were

Table 4. Correlation between Age and Genotype at c.4603C→T (*rs1676486*) in *COL11A1*

Population	Mean ± SD Age (in years) for Genotype			*P*[a]
	CC	CT	TT	
Case	36.8 ± 15.0	36.9 ± 14.5	36.8 ± 14.5	.58
Control	64.8 ± 12.1	63.9 ± 11.1	63.1 ± 13.1	.54

[a] *P* value was calculated using the Kruskal-Wallis test.

Table 5. Genotype at c.4603C→T (*rs1676486*) in *COL11A1*, Stratified by Sex

Measure	Male			Female		
	Case	Control	Total	Case	Control	Total
No. of subjects:						
All	298	177	475	191	285	476
CC	116	98	214	88	154	242
CT	144	65	209	79	112	191
TT	38	14	52	24	19	43
T allele frequency (%)	.37	.26	.33	.33	.26	.29
P value[a]	...	...	.00074	...	...	.021

[a] *P* value for allelic difference between the patients with LDH and the control groups for each sex, by the χ^2 test.

digested using *Hin*dIII, and *COL11A1* mRNAs were transcribed using RiboMax Large Scale RNA Production System-T7 (Promega) and were purified by SV Total RNA Isolation System (Promega). The whole-cell extract was prepared by washing OUMS-27 cells in PBS and resuspending them in an extraction buffer. After incubation on ice for 30 min and microcentrifugation at 4°C, we transferred supernatants to new tubes and stored them at −80°C until use. We mixed and incubated each 5 μg of synthesized RNA and the diluted (1:1,000) whole-cell extract at room temperature for the tested time (5 or 10 min). We stopped the reaction with addition of a formamide dye. The samples were then heated at 95°C for 5 min and were placed on ice immediately. We detected *COL11A1* mRNAs of the samples by northern-blot analysis and quantified their signal intensities, using the Esper-Scanner (Epson) and Adobe Photoshop 6.0.

Immunohistochemistry for Type XI Collagen

We processed and embedded tissue samples in paraffin by the AMeX method. We predigested the tissue sections with 500 U/ml of testicular hyaluronidase (Sigma) for 30 min at 37°C. For immunofluorescent visualization, we blocked nonspecific labeling with blocking reagent (DakoCytomation) for 10 min at room temperature and then incubated the sections with the rabbit polyclonal antibody against bovine type XI collagen (1:500) at 4°C overnight. For the staining of the negative control, we applied nonimmune rabbit IgG (DakoCytomation) to the section instead of primary antibody. After washing them with Tris-buffered saline, we incubated the sections with secondary antibody conjugated to horseradish peroxidase–labeled polymer (Envision+ [DakoCytomation]) for 30 min at room temperature. We visualized the immunoreactive products using a diaminobenzidine reagent and counterstained them with hematoxylin.

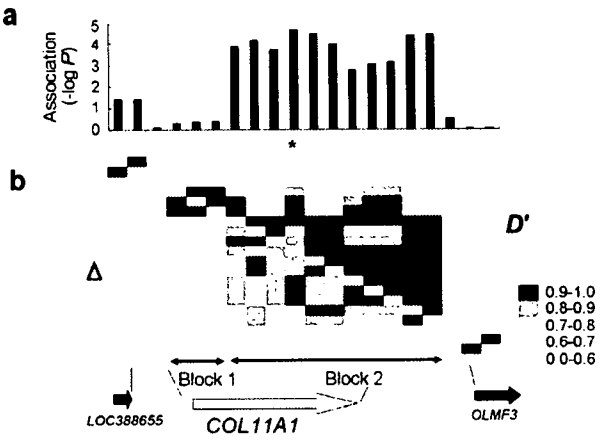

Figure 1. Case-control association study and linkage-disequilibrium mapping. *a*, Association of *COL11A1* with LDH. The $-\log_{10}$ transformation of the corrected *P* value (allele 1 vs. allele 2) was plotted on the *Y*-axis. The asterisk (*) indicates c.4603C→T. *b*, Pairwise linkage disequilibrium between SNPs in and around *COL11A1* measured by *D'* and Δ in 465 controls. The *COL11A1* region is divided into two linkage-disequilibrium blocks.

Results

We first examined the association of the type XI collagen genes (*COL11A1*, *COL11A2*, and *COL2A1*) with LDD, which included patients with and without LDH. We tested tag SNPs that were selected from the JSNP Database and the International HapMap Project database. A comparison of 188 LDD cases and 179 controls revealed no association with any of the SNPs; however, there was a significant association with *COL11A1* when cases were stratified on the basis of the presence or absence of LDH (table 2). In a comparison of 130 patients with LDH with 179 controls, one SNP (c.4603C→T [*rs1676486*]) had a significant association. To confirm the association, we examined another 359 LDH cases and 286 controls for the *COL11A1* SNP. Again, we identified the significant association between the SNP and LDH (table 2). Adjusted *P* = .00030 was obtained by 10^7 permutations

To identify the disease-causing sequence variation, we examined sequence variations in *COL11A1* exons and their flanking regions from a public database and by resequencing 24 patients with LDH. A total of 23 sequence variations were identified and were tested for association. SNP c.4603C→T had the most significant association (table 3), which remained significant after Bonferroni correction for multiple testing. We examined whether confounding effects, such as age and sex, affect the associations with LDH and found no relationship between the genotype and

Table 6. Haplotype Association Analysis of *COL11A1* with LDH

Haplotype	Frequency		*P*[a]
	Case	Control	
H1	.527	.616	.000154
H2	.302	.222	.000150
H3	.038	.039	.90
H4	.041	.037	.63
H5	.045	.034	.27
H6	.014	.014	.91
H7	.011	.008	.50

NOTE.—Results are for the haplotypes of block 2 that contained the susceptibility SNP, c.4603C→T.

[a] By the χ^2 test.

Table 7. Association between Genotype at c.4603C→T (*rs1676486*) in *COL11A1* and LDH in the Japanese Population

Group	No. with Genotype			Allelic Frequency	*P*	OR (95% CI)
	CC	CT	TT			
Case	360	367	96	.34	.0000033	1.42 (1.23–1.65)
Control	453	325	60	.265		

these factors (table 4). The association was positive in both sexes (table 5).

Using the 20 SNPs in and around *COL11A1* that had a minor-allele frequency >10%, we analyzed the linkage-disequilibrium structure of the region and found highly structured linkage-disequilibrium blocks (fig. 1). *COL11A1* was covered by two blocks, and the SNP with a significant association (c.4603C→T) was contained in block 2. We further analyzed the haplotype structure of block 2 and identified seven haplotypes with frequencies >0.01 that covered >97% of both the case and control groups (table 6). The association was weaker than that of c.4603C→T alone, suggesting the absence of a hidden causal SNP. We further examined the association of the SNP, using an additional 334 patients with LDH and 376 controls. Our findings of the association between this SNP and LDH were replicated (P = .044; OR 1.27 [95% CI 1.01–1.59]. Therefore, this SNP is strongly associated with LDH (combined $P = 3.3 \times 10^{-6}$ in allelic frequency) (table 7).

To clarify the functional impact of c.4603C→T, we quantified the allelic difference of the mRNA expression by real-time RT-PCR. The expression level of the susceptibility allele c.4603T was significantly lower than that of the c.4603C allele (fig. 2*a*). We hypothesized that this SNP affects *COL11A1* transcription by altering mRNA stability and examined the stability of *COL11A1* mRNA containing the SNP. We mixed mRNAs produced by in vitro transcription with cell lysate and assessed mRNA degradation by endogenous components of the cells, using northern-blot analysis. The transcript containing the susceptible allele degraded faster (fig. 2*b* and 2*c*).

To gain insight into the role of type XI collagen in LDH, we examined *COL11A1* expression in tissues and cells by quantitative real-time PCR. *COL11A1* mRNA was predominantly expressed in IVD (fig. 3*a*). We investigated the correlation between the *COL11A1* mRNA expression level and a variety of LDH phenotypes and found that severity of disc degeneration evaluated by MRI was inversely correlated with *COL11A1* expression in IVDs of patients with LDH (fig. 3*b*). We further analyzed the expression and localization of type XI collagen in IVD by immunohistochemistry. Normal discs had a highly uniform ECM structure, with intense immunostaining of type XI collagen in the nucleus pulposus cells and ECM (fig. 3*c*). In degenerative discs, however, we observed weak immunostaining of type XI collagen around the nucleus pulposus cells (fig. 3*d*). These findings implicate a decrease of type XI collagen in the pathogenesis of LDH.

Discussion

Through a case-control association study focusing on type XI collagen, we identified *COL11A1* as a susceptibility gene for LDH. *COL11A1* mRNA was substantially ex-

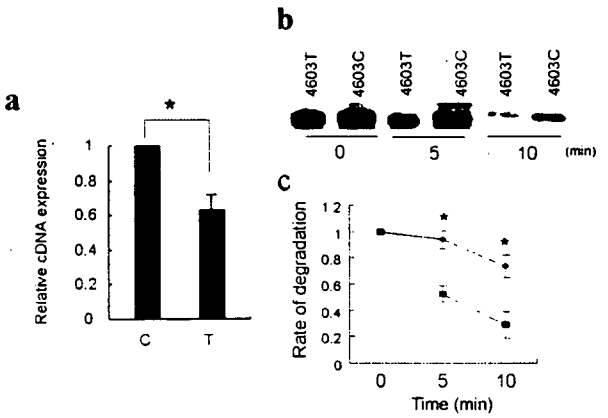

Figure 2. Difference in transcription and stability of *COL11A1* mRNA containing the LDH-associated SNP. *a*, Relative cDNA expression of c.4603C→T evaluated by real-time PCR. Data represent the ratios of cDNA to genomic DNA, and expression of the C allele is converted to 1 (an asterisk [*] indicates $P < .05$, by Student's t test). Data represent the mean ± SD in triplicate assays. *b*, Sequential change of *COL11A1* mRNA analyzed by northern blotting. "4603C" and "4603T" indicate *COL11A1* mRNA produced by in vitro transcription with c.4603C and c.4603T, respectively. *c*, Rate of degradation of the transcripts. Diamonds indicate the transcript with c.4603C; squares indicate the transcript with c.4603T. The difference of the rate of degradation was significant at both 5 min and 10 min after the reaction (an asterisk (*) indicates $P < .05$, by Student's t test). Data represent the mean ± SD in triplicate assays.

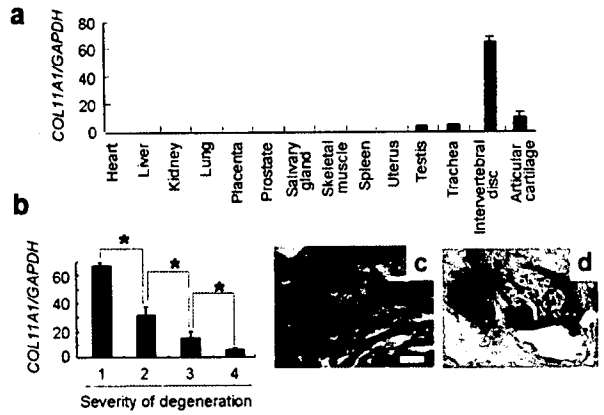

Figure 3. Type XI collagen expression in human. *a*, *COL11A1* expression in different tissues. *COL11A1* mRNA was predominantly expressed in IVD. *b*, Inverse correlation between *COL11A1* expression and severity of degeneration of IVD in patients with LDH (an asterisk [*] indicates $P < .05$, by Student's *t* test). The degree of disc degeneration is evaluated by MRI and is scored according to the classification of Schneiderman. *c* and *d*, Immunostaining of type XI collagen in IVDs from an unaffected individual (*c*) and a patient with LDH (Schneiderman's grade 3) (*d*). Ubiquitous and intense staining was found in the normal disc. In contrast, the staining was found only in and around the territorial matrices of clustered cells in the degenerative disc. The white scale bar indicates 50 nm.

pressed in IVD, and the expression in patients with LDH was decreased according to the severity of degeneration. Our findings further indicate that the susceptibility SNP produces unstable *COL11A1* transcripts. A few *cis*-elements have been implicated in mRNA stabilization.[25] The 4856–4865 nucleotides (caaaaaatct) in *COL11A1* mRNA closely match the consensus for a mRNA stability motif, "g/tanaaaag/tcc/t."[26] The sequence variation might affect the mRNA stability motif and disrupt the *cis*-element critical for mRNA stability, although they are >200 bp apart. Alternatively, the sequence variation might induce a conformational change in the mRNA that would decrease mRNA stability or increase the sensitivity to RNase. The decrease of the *COL11A1* transcript would lead to a decrease in type XI collagen in the ECM of IVD.

IVD has a highly structured ECM to resist mechanical forces. The highly oriented network of the fibrillar collagens offers tensile strength,[27,28] and highly hydrated aggregating PG resists comprehensive forces. They form a mesh suited to holding water molecules, which further increases their ability to withstand mechanical forces. Therefore, the structural integrity of ECM and the physiologic balance of its components are critical to IVD function. Perturbation of ECM metabolism would increase the mechanical load of the IVD, leading to its degeneration. The reduction in type XI collagen, the critical organizer of ECM, ultimately causes disintegration of ECM and hence IVD degeneration, although it could occur as a secondary event of LDH. The present study underscores the importance of ECM proteins in the pathogenesis of common bone and joint diseases, including LDH. Our results should lead to a better understanding of the pathogenic mechanisms of LDH and suggest promising targets for a novel treatment strategy for LDH.

Acknowledgments

We thank Drs. S. Seki, T. Koyanagi, Y. Fukui, T. Kono, H. Kono, H. Hirabayashi, K. Kono, M. Ishikawa, M. Tamura, K. Nojiri, H. Morisue, and N. Hosogane, for help in collecting samples and performing the experimental study, and Y. Takanashi, for technical assistance. This work was supported by grants-in-aid from Ministry of Education, Culture, Sports and Science of Japan (grant 19209049 [to S.I.]).

Web Resources

Accession numbers and URLs for data presented herein are as follows:

Applied Biosystems, http://www.appliedbiosystems.com/index.cfm

GenBank, http://www.ncbi.nlm.nih.gov/Genbank/ (for *COL11A1* sequences [accession numbers NM001854.2, AC093150.4, AL627203.7, and AC099567.2])

International HapMap Project, http://hapmap.org/

JSNP Database, http://snp.ims.u-tokyo.ac.jp/index.html

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for Stickler syndrome type II, Marshall syndrome, and oto-spondylo-mega-epiphyseal dysplasia)

References

1. Andersson GB (1999) Epidemiological features of chronic low-back pain. Lancet 354:581–585
2. Skovron ML (1992) Epidemiology of low back pain. Baillieres Clin Rheumatol 6:559–573
3. Videman T, Battie MC (1996) A critical review of epidemiology of idiopathic low back pain. In: Weinstein JN, Gordon SL (eds) Low back pain: a scientific and clinical overview. American Academy of Orthopaedic Surgeons, Rosemont, IL, pp 317–332
4. Matsui H, Kanamori M, Ishihara H, Yudoh K, Naruse Y, Tsuji

H (1998) Familial predisposition for lumbar degenerative disc disease: a case-control study. Spine 23:1029–1034
5. Sambrook PN, MacGregor AJ, Spector TD (1999) Genetic influences on cervical and lumbar disc degeneration: a magnetic resonance imaging study in twins. Arthritis Rheum 42: 366–372
6. Kimura T, Nakata K, Tsumaki N, Miyamoto S, Matsui Y, Ebara S, Ochi T (1996) Progressive degeneration of articular cartilage and intervertebral discs: an experimental study in transgenic mice bearing a type IX collagen mutation. Int Orthop 20: 177–181
7. Watanabe H, Nakata K, Kimata K, Nakanishi I, Yamada Y (1997) Dwarfism and age-associated spinal degeneration of heterozygote cmd mice defective in aggrecan. Proc Natl Acad Sci USA 94:6943–6947
8. Annunen S, Paassilta P, Lohiniva J, Perala M, Pihlajamaa T, Karppinen J, Tervonen O, Kroger H, Lahde S, Vanharanta H, et al (1999) An allele of *COL9A2* associated with intervertebral disc disease. Science 285:409–412
9. Paassilta P, Lohiniva J, Goring HH, Perala M, Raina SS, Karppinen J, Hakala M, Palm T, Kroger H, Kaitila I, et al (2001) Identification of a novel common genetic risk factor for lumbar disk disease. JAMA 285:1843–1849
10. Kawaguchi Y, Osada R, Kanamori M, Ishihara H, Ohmori K, Matsui H, Kimura T (1999) Association between an aggrecan gene polymorphism and lumbar disc degeneration. Spine 24: 2456–2460
11. Seki S, Kawaguchi Y, Chiba K, Mikami Y, Kizawa H, Oya T, Mio F, Mori M, Miyamoto Y, Masuda I, et al (2005) A functional SNP in *CILP*, encoding cartilage intermediate layer protein, is associated with susceptility to intervertebral disc disease of the lumbar spine. Nat Genet 37:607–612
12. Kizawa H, Kou I, Iida A, Sudo A, Miyamoto Y, Fukuda A, Mabuchi A, Kotani A, Kawakami A, Yamamoto S, et al (2005) An asparatic acid repeat polymorphism in asporin negatively affects chondrogenesis and increase susceptibility to osteoarthritis. Nat Genet 37:138–144
13. Keene DR, Oxford JT, Morris NP (1995) Ultrastructural localization of collagen types II, IX, and XI in the growth plate of human rib and fetal bovine epiphyseal cartilage: type XI collagen is restricted to thin fibrils. J Histochem Cytochem 43:967–979
14. Blaschke UK, Eikenberry EF, Hulmes DJ, Galla HJ, Bruckner P (2000) Collagen XI nucleates self-assembly and limits lateral growth of cartilage fibrils. J Biol Chem 275:10370–10378
15. Gregory KE, Oxford JT, Chen Y, Gambee JE, Gygi SP, Aebersold R, Neame PJ, Mechling DE, Bachinger HP, Morris NP (2000) Structural organization of distinct domains within the non-collagenous N-terminal region of collagen type XI. J Biol Chem 275:11498–11506
16. Li Y, Lacerda DA, Warman ML, Beier DR, Yoshioka H, Ninomiya Y, Oxford JT, Morris NP, Andrikopoulos K, Ramirez F, et al (1995) A fibrillar collagen gene, *Col11a1*, is essential for skeletal morphogenesis. Cell 80:423–430
17. Seegmiller RE, Myers RA, Dorfman A, Horwitz AL (1981) Structural and associative properties of cartilage matrix constituents in mice with hereditary chondrodysplasia. Connect Tissue Res 9:69–77
18. Oegema TR Jr (1993) Biochemistry of the intervertebral disc. Clin Sports Med 12:419–439
19. Smith GN Jr, Williams JM, Brandt KD (1985) Interaction of proteoglycans with the pericellular (1α, 2α, 3α) collagens of cartilage. J Biol Chem 260:10761–10767
20. Spranger J (1998) The type XI collagenopathies. Pediatr Radiol 28:745–750
21. Rose PS (2001) Thoracolumbar spinal abnormalities in Stickler syndrome. Spine 26:403–409
22. Schneiderman G, Flannigan B, Kingston S, Thomas J, Dillin WH, Watkins RG (1987) Magnetic resonance imaging in the diagnosis of disc degeneration: correlation with discography. Spine 12:276–281
23. Tanaka T, Ikari K, Furushima K, Okada A, Tanaka H, Furukawa K, Yoshida K, Ikeda T, Ikegawa S, Hunt SC, et al (2003) Genomewide linkage and linkage disequilibrium analyses identify *COL6A1*, on chromosome 21, as the locus for ossification of the posterior longitudinal ligament of the spine. Am J Hum Genet 73:812–822
24. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, et al (2007) A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 445:881–885
25. Mullner EW, Kuhn LC (1988) A stem-loop in the 3′ untranslated region mediates iron-dependent regulation of transferrin receptor mRNA stability in the cytoplasm. Cell 53:815–825
26. Capon F, Allen MH, Ameen M, Burden AD, Tillman D, Barker JN, Trembath RC (2004) A synonymous SNP of the corneodesmosin gene leads to increased mRNA stability and demonstrates association with psoriasis across diverse ethnic groups. Hum Mol Genet 13:2361–2368
27. Kempson GE, Muir H, Pollard C, Tuke M (1973) The tensile properties of cartilage of human femoral condyles related to the content of collagen and glycosaminoglycans. Biochem Biophys Acta 297:456–472
28. Schmidt MD, Mow VC, Chun LE, Eyre DR (1990) Effects of proteoglycan extraction on the tensile behavior of articular cartilage. J Orthop Res 8:353–363

LETTERS

Genome-wide detection and characterization of positive selection in human populations

Pardis C. Sabeti[1]*, Patrick Varilly[1]*, Ben Fry[1], Jason Lohmueller[1], Elizabeth Hostetter[1], Chris Cotsapas[1,2], Xiaohui Xie[1], Elizabeth H. Byrne[1], Steven A. McCarroll[1,2], Rachelle Gaudet[3], Stephen F. Schaffner[1], Eric S. Lander[1,4,5,6] & The International HapMap Consortium†

With the advent of dense maps of human genetic variation, it is now possible to detect positive natural selection across the human genome. Here we report an analysis of over 3 million polymorphisms from the International HapMap Project Phase 2 (HapMap2)[1]. We used 'long-range haplotype' methods, which were developed to identify alleles segregating in a population that have undergone recent selection[2], and we also developed new methods that are based on cross-population comparisons to discover alleles that have swept to near-fixation within a population. The analysis reveals more than 300 strong candidate regions. Focusing on the strongest 22 regions, we develop a heuristic for scrutinizing these regions to identify candidate targets of selection. In a complementary analysis, we identify 26 non-synonymous, coding, single nucleotide polymorphisms showing regional evidence of positive selection. Examination of these candidates highlights three cases in which two genes in a common biological process have apparently undergone positive selection in the same population: *LARGE* and *DMD*, both related to infection by the Lassa virus[3], in West Africa; *SLC24A5* and *SLC45A2*, both involved in skin pigmentation[4,5], in Europe; and *EDAR* and *EDA2R*, both involved in development of hair follicles[6], in Asia.

An increasing amount of information about genetic variation, together with new analytical methods, is making it possible to explore the recent evolutionary history of the human population. The first phase of the International Haplotype Map, including ~1 million single nucleotide polymorphisms (SNPs)[7], allowed preliminary examination of natural selection in humans. Now, with the publication of the Phase 2 map (HapMap2)[1] in a companion paper, over 3 million SNPs have been genotyped in 420 chromosomes from three continents (120 European (CEU), 120 African (YRI) and 180 Asian from Japan and China (JPT + CHB)).

In our analysis of HapMap2, we first implemented two widely used tests that detect recent positive selection by finding common alleles carried on unusually long haplotypes[2]. The two, the Long-Range Haplotype (LRH)[8] and the integrated Haplotype Score (iHS)[9] tests, rely on the principle that, under positive selection, an allele may rise to high frequency rapidly enough that long-range association with nearby polymorphisms—the long-range haplotype[8]—will not have time to be eliminated by recombination. These tests control for local variation in recombination rates by comparing long haplotypes to other alleles at the same locus. As a result, they lose power as selected alleles approach fixation (100% frequency), because there are then few alternative alleles in the population (Supplementary Fig. 2 and Supplementary Tables 1–2).

We next developed, evaluated and applied a new test, Cross Population Extended Haplotype Homozogysity (XP-EHH), to detect selective sweeps in which the selected allele has approached or achieved fixation in one population but remains polymorphic in the human population as a whole (Methods, and Supplementary Fig. 2 and Supplementary Tables 3–6). Related methods have recently also been described[10–12].

Our analysis of recent positive selection, using the three methods, reveals more than 300 candidate regions[1](Supplementary Fig. 3 and Supplementary Table 7), 22 of which are above a threshold such that no similar events were found in 10 Gb of simulated neutrally evolving sequence (Methods). We focused on these 22 strongest signals (Table 1), which include two well-established cases, *SLC24A5* and *LCT*[2,5,13], and 20 other regions with signals of similar strength.

The challenge is to sift through genetic variation in the candidate regions to identify the variants that were the targets of selection. Our candidate regions are large (mean length, 815 kb; maximum length, 3.5 Mb) and often contain multiple genes (median, 4; maximum, 15). A typical region harbours ~400–4,000 common SNPs (minor allele frequency >5%), of which roughly three-quarters are represented in current SNP databases and half were genotyped as part of HapMap2 (Supplementary Table 8).

We developed three criteria to help highlight potential targets of selection (Supplementary Fig. 1): (1) selected alleles detectable by our tests are likely to be derived (newly arisen), because long-haplotype tests have little power to detect selection on standing (pre-existing) variation[14]; we therefore focused on derived alleles, as identified by comparison to primate outgroups; (2) selected alleles are likely to be highly differentiated between populations, because recent selection is probably a local environmental adaptation[2]; we thus looked for alleles common in only the population(s) under selection; (3) selected alleles must have biological effects. On the basis of current knowledge, we therefore focused on non-synonymous coding SNPs and SNPs in evolutionarily conserved sequences. These criteria are intended as heuristics, not absolute requirements. Some targets of selection may not satisfy them, and some will not be in current SNP databases. Nonetheless, with ~50% of common SNPs in these populations genotyped in HapMap2, a search for causal variants is timely.

We applied the criteria to the regions containing *SLC24A5* and *LCT*, each of which already has a strong candidate gene, mutation and trait. At *SLC24A5*, the 600 kb region contains 914 genotyped

[1]Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02139, USA. [2]Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [3]Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138, USA. [4]Department of Biology, MIT, Cambridge, Massachusetts 02139, USA. [5]Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, USA. [6]Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA
*These authors contributed equally to this work.
†Lists of participants and affiliations appear at the end of the paper.

Table 1 | The twenty-two strongest candidates for natural selection

Region	Chr:position (MB, HG17)	Selected population	Long Haplotype Test	Size (Mb)	Total SNPs with Long Haplotype Signal	Subset of SNPs that fulfil criteria 1	Subset of SNPs that fulfil criteria 1 and 2	Subset of SNPs that fulfil criteria 1, 2 and 3	Genes at or near SNPs that fulfil all three criteria
1	chr1:166	CHB + JPT	LRH, iHS	0.4	92	39	30	2	*BLZF1, SLC19A2*
2	chr2:72.6	CHB + JPT	XP-EHH	0.8	732	250	0	0	
3	chr2:108.7	CHB + JPT	LRH, iHS, XP-EHH	1.0	972	265	7	1	*EDAR*
4	chr2:136.1	CEU	LRH, iHS, XP-EHH	2.4	1,213	282	24	3	*RAB3GAP1, R3HDM1, LCT*
5	chr2:177.9	CEU,CHB + JPT	LRH, iHS, XP-EHH	1.2	1,388	399	79	9	*PDE11A*
6	chr4:33.9	CEU,YRI, CHB + JPT	LRH, iHS	1.7	413	161	33	0	
7	chr4:42	CHB + JPT	LRH, iHS, XP-EHH	0.3	249	94	65	6	*SLC30A9*
8	chr4:159	CHB + JPT	LRH, iHS, XP-EHH	0.3	233	67	34	1	
9	chr10:3	CEU	LRH, iHS, XP-EHH	0.3	179	63	16	1	
10	chr10:22.7	CEU, CHB + JPT	XP-EHH	0.3	254	93	0	0	
11	chr10:55.7	CHB + JPT	LRH, iHS, XP-EHH	0.4	735	221	5	2	*PCDH15*
12	chr12:78.3	YRI	LRH, iHS	0.8	151	91	25	0	
13	chr15:46.4	CEU	XP-EHH	0.6	867	233	5	1	*SLC24A5*
14	chr15:61.8	CHB + JPT	XP-EHH	0.2	252	73	40	6	*HERC1*
15	chr16:64.3	CHB + JPT	XP-EHH	0.4	484	137	2	0	
16	chr16:74.3	CHB + JPT, YRI	LRH, iHS	0.6	55	35	28	3	*CHST5, ADAT1, KARS*
17	chr17:53.3	CHB + JPT	XP-EHH	0.2	143	41	0	0	
18	chr17:56.4	CEU	XP-EHH	0.4	290	98	26	3	*BCAS3*
19	chr19:43.5	YRI	LRH, iHS, XP-EHH	0.3	83	30	0	0	
20	chr22:32.5	YRI	LRH	0.4	318	188	35	3	*LARGE*
21	chr23:35.1	YRI	LRH, iHS	0.6	50	35	25	0	
22	chr23:63.5	YRI	LRH, iHS	3.5	13	3	1	0	
		Total SNPs		16.74	9,166	2,898	480	41	

Twenty-two regions were identified at a high threshold for significance (Methods), based on the LRH, iHS and/or XP-EHH test. Within these regions, we examined SNPs with the best evidence of being the target of selection on the basis of having a long haplotype signal, and by fulfilling three criteria: (1) being a high-frequency derived allele; (2) being differentiated between populations and common only in the selected population; and (3) being identified as functional by current annotation. Several candidate polymorphisms arise from the analysis including well-known *LCT* and *SLC24A5* (ref. 2), as well as intriguing new candidates.

SNPs. Applying filters progressively (Table 1 and Fig. 1a–d), we found that 867 SNPs are associated with the long-haplotype signal, of which 233 are high-frequency derived alleles, of which 12 are highly differentiated between populations, and of which only 5 are common in Europe and rare in Asia and Africa. Among these five SNPs, there is only one implicated as functional by current knowledge; it has the strongest signal of positive selection and encodes the A111T polymorphism associated with pigment differences in

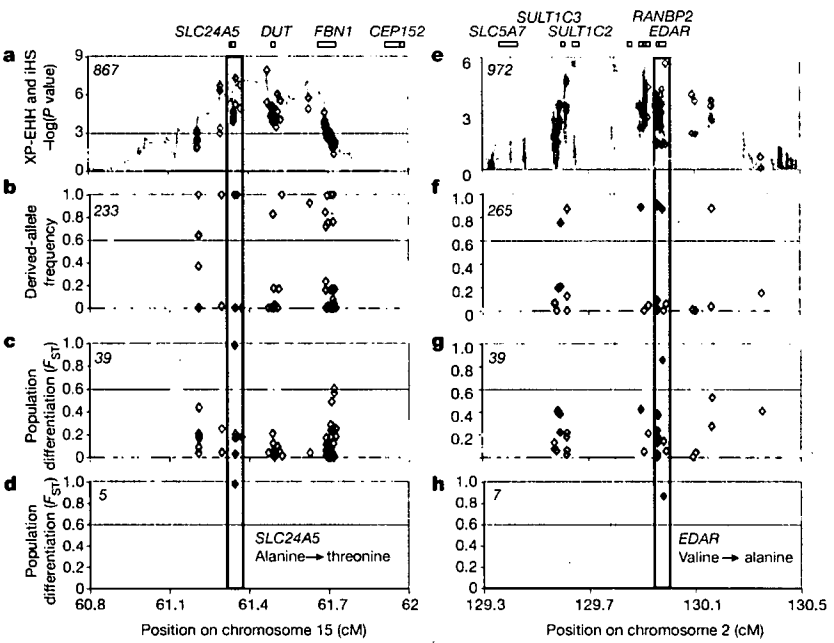

Figure 1 | Localizing *SLC24A5* and *EDAR* signals of selection. **a–d**, *SLC24A5*. **a**, Strong evidence for positive selection in CEU samples at a chromosome 15 locus: XP-EHH between CEU and JPT + CHB (blue), CEU and YRI (red), and YRI and JPT + CHB (grey). SNPs are classified as having low probability (bordered diamonds) and high probability (filled diamonds) potential for function. SNPs were filtered to identify likely targets of selection on the basis of the frequency of derived alleles (**b**), differences between populations (**c**) and differences between populations for high-frequency derived alleles (less than 20% in non-selected populations) (**d**). The number of SNPs that passed each filter is given in the top left corner in red. The threonine to alanine candidate polymorphism in *SLC24A5* is the clear outlier. **e–h**, *EDAR*. **e**, Similar evidence for positive selection in JPT + CHB at a chromosome 2 locus: XP-EHH between CEU and JPT + CHB (blue), between YRI and JPT + CHB (red), and between CEU and YRI (grey); iHS in JPT + CHB (green). A valine to alanine polymorphism in *EDAR* passes all filters: the frequency of derived alleles (**f**), differences between populations (**g**) and differences between populations for high-frequency derived alleles (less than 20% in non-selected populations) (**h**). Three other functional changes, a D→E change in *SULT1C2* and two SNPs associated with *RANBP2* expression (Methods), have also become common in the selected population.