SCIENTIFIC REPORTS

OPEN

SUBJECT AREAS:
GENOME-WIDE ASSOCIATION STUDIES
RNA SPLICING
SKIN MODELS
HAPLOTYPES

Received
17 December 2012

Accepted
22 May 2013

Published
17 June 2013

Correspondence and requests for materials should be addressed to B.G. (bgandolfi@ucdavis.edu)

A splice variant in *KRT71* is associated with curly coat phenotype of Selkirk Rex cats

Barbara Gandolfi[1], Hasan Alhaddad[1], Shannon E. K. Joslin[1], Razib Khan[1], Serina Filler[2], Gottfried Brem[2] & Leslie A. Lyons[1]

[1]Department of Population Health and Reproduction, School of Veterinary Medicine, University of California - Davis, Davis, CA USA, [2]Institute of Animal Breeding and Genetics, Department for Biomedical Sciences, University of Veterinary Medicine - Vienna, Vienna, Austria.

One of the salient features of the domestic cat is the aesthetics of its fur. The Selkirk Rex breed is defined by an autosomal dominant woolly rexoid hair (ADWH) abnormality that is characterized by tightly curled hair shafts. A genome-wide case – control association study was conducted using 9 curly coated Selkirk Rex and 29 controls, including straight-coated Selkirk Rex, British Shorthair and Persian, to localize the Selkirk autosomal dominant rexoid locus (SADRE). Although the control cats were from different breed lineages, they share recent breeding histories and were validated as controls by Bayesian clustering, multi-dimensional scaling and genomic inflation. A significant association was found on cat chromosome B4 ($P_{raw} = 2.87 \times 10^{-11}$), and a unique haplotype spanning ~600 Kb was found in all the curly coated cats. Direct sequencing of four candidate genes revealed a splice site variant within the *KRT71* gene associated with the hair abnormality in Selkirk Rex.

Body homeostasis and tissue integration are supported by the hair, a highly keratinized tissue produced in the hair follicle (HF). Hair formation in mammalian HFs occurs during embryogenesis through a series of reciprocal interactions between skin epithelial and underlying dermal cells[1,2]. The HF undergoes dynamic cell kinetics composed of the anagen (active growth) phase, the catagen (transition) phase, and the telogen (resting) phase[3]. Among the skin appendages, the HF has a highly complex structure with eight distinct cell layers, where hundreds of gene products play key roles in hair cycle and maintenance. The hair shaft is surrounded and supported by the inner root sheath, the companion layer, and the outer root sheath[4]. The formation of a rigid structure during the HF differentiation is due to keratin proteins that are abundantly and differentially expressed[2,4,5].

A mammal's pelage is generally one of its most noticeable attributes and is aesthetically pleasing. In cats, coat color and pelage types are often selected as a specific trait to develop a breed. The coat of a normal cat consists three hair types: long and straight guard hairs of uniform diameter, thinner awn hairs, and the fine undulating down hairs of uniform thickness[6]. Rexoid (curly / woolly) pelage is an easily recognized trans-species anomaly; detailed studies in various mammalian species, including mice[7,8], chicken[9], rat[10], dog[11] and human[12–15], have identified causative genes and mutations. Nine rexoid-type pelage phenotypes are known within the domestic cat[16–20] and recently, significant advances have been made toward the identification of these feline hair abnormalities. One study revealed two alleles ($KRT71^{re}$ and $KRT71^{hr}$) within *KRT71*, a crucial gene for keratinization in the HF, which are responsible for recessive hypotricosis *Hairless* (*Hr, hr*) locus of the Sphynx breed and the *Rex* hair locus (*Re, re*) of the curly coated Devon Rex breed[19]. A second rexoid locus (*R, r*) with a mutation within *P2RY5* is responsible for the autosomal recessive woolly hair in the Cornish Rex breed (Gandolfi 2013, in press). For each of these autosomal recessive rexoid / woolly hair conditions, the identified mutations are responsible for a major change in the hair follicle, altering hair formation. Several dominant rexoid / woolly hair conditions define other breeds, such as Selkirk Rex, LaPerm and American Wirehair, and these mutations await identification and characterization.

The Selkirk Rex breed harbors the most recently derived rexoid coat mutation that has successfully been developed into an internationally recognized breed in the domestic cat fancy. The breed is proposed to have originated from a *de novo* mutation in a domestic cat in 1987, and a previous study estimates that an average of only 8.4 generations elapsed since the occurrence of the Selkirk Rex rexoid mutation[21]. To maintain genetic diversity and import aesthetic features, such as a round and sturdy body type, wide-set eyes and ears, long hair and a moderately brachycephalic muzzle, the Selkirk Rex has been extensively crossed to British Shorthair, Exotic Shorthair, and Persians to establish the breed. Filler *et al.* (2012) confirmed the out-crossing history, and shows that British Shorthair, Exotic Shorthair, Persian, Selkirk Rex, and Scottish Fold constitute breed lineages of a higher order grouping, referred to as the "Persian family" of breeds.

The autosomal dominant forms of woolly hair / hair loss disorders in human have been shown to result from mutations in *corneodesmosin* (*CDSN*)[22], the inhibitory upstream open reading frame (U2HR) of *hairless* (*HR*)[23,24], *APC-downregulated-by-1* (*APCDD1*)[25], *Kertain-71*[26] and *Keratin-74*[14]. Yet human studies have not fully elucidated the molecular basis of hereditary hair disorders; hence, a better determination of pathogenic dominant mutations within genes is necessary. The rexoid domestic cat breeds are a model that can reveal new loci involved in hair texture and hair shaft anomalies, as well as characterize the effect of different mutations within the same loci as found in humans and other species.

A genome-wide association study (GWAS) was performed, using the illumina Infinium feline 63 K iSelect DNA array, on curly and straight Selkirk Rex cats to identify the Selkirk Autosomal Dominant Rex (SADRE)[21] locus responsible for the autosomal dominant rexoid hair in this cat breed. The dominant nature of the Selkirk Rex curly coat trait should provide an ideal situation for a case-control genome-wide association study. However, due to the lack of sufficient number of straight hair Selkirk Rex individuals, Persian and Scottish Fold were used as controls, as these breeds possess the same population genetic background as Selkirk Rex.

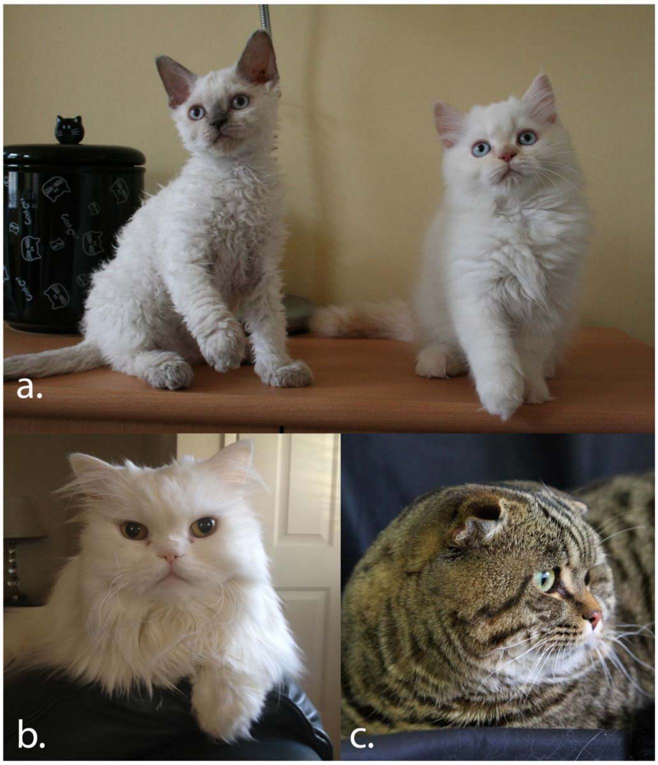

Figure 1 | Phenotypic variation within the Persian-family of cat breeds. (a) Phenotype of the Selkirk Rex breed characterized by a curly coated cats (left) as well as straight hair cat (right) and moderate brachycephaly. (b) Phenotype of the Persian breed characterized by straight long hair and moderate brachycephaly. (c) Phenotype of Scottish Fold breed characterized by folded ears and moderate brachycephaly. Photos by Barbara Gandolfi.

Results

Genome-wide association study. A dataset composed of 17 curly coated Selkirk Rex, and 62 straight haired cats (Selkirk Rex (n = 7), Scottish Fold (n = 32), and Persian (n = 23)) (**Figure 1**) were genotyped on the DNA array. Only one Scottish Fold was excluded due to low genotyping rate (MIND > 0.2). The remaining samples were evaluated for population stratification and relatedness. The stratification analysis assessed by multi-dimensional scaling (MDS) revealed 3 main clusters (**Figure S1a**). One cluster consisted of a majority of the Persians; a second cluster containing the Selkirk Rex cases and controls, and a third cluster represented by a small subset of Scottish Folds. The remainder of the Persians, and the majority of the Scottish Folds, spanned the interval across the two antipodal populations. The two Persian and Scottish Fold outlier groups (n = 18 and n = 7, respectively) were excluded from further analysis. After removing these outliers, 17 cases and 36 controls (n = 53) were evaluated for relatedness. Nine Selkirk Rex individuals (one with straight hair and eight with curly hair) and six Scottish Fold exhibited a $\widehat{P} > 0.3$ and were excluded from the dataset for the association study (**Figure S1b**). The exclusion of the related individuals reduced the genomic inflation factor (λ) from 2.15 to 1.46 and is depicted on a quantile-quantile plot (Q-Q) (**Figure S2**).

A genome-wide association analysis of nine cases (curly coated Selkirk Rex) and 29 controls (six straight Selkirk Rex, five Persian and 18 Scottish Fold) was performed. After genotyping pruning of the 62,897 array SNPs for low minor allele frequency (MAF) and poor SNP genotypes, 10,344 SNPs were removed from further analysis due to a MAF $< 5\%$ and 677 SNPs failed missingness (GENO > 0.2). The highest significant association was identified on cat chromosome B4 at position 81,264,280 (Felis_cats 6.2/felCat5 release) (**Figure 2**). The *P* values, chromosomal location and SNP frequency of the 10 most associated SNPs are presented in **Table 1**. After performing permutations to correct P_{raw} values for multiple hypothesis testing, 13 SNPs on chromosome B4 reached genome wide significance ($P_{genome} = 0.00058$ for SNP B4.96694363 at position 81,264,280), as well as one less significant SNP on chromosome B3 ($P_{genome} = 0.02$) and 2 on chromosome "unknown", which contains the yet unassembled contigs ($P_{genome} = 0.019$ for both SNPs) (**Table S1**). The strongest association was detected on chromosome B4 between positions 81,264,280 to 81,980,475. To further support the association, a genome-wide F_{st} analysis was performed between unrelated curly Selkirk Rex and each of the control populations, separately and combined (**Figure S3**). The region that differentiated between the curly Selkirk Rex and any of the control populations overlapped with the association region, particularly as demonstrated by the Selkirk Rex and Scottish Fold comparison (**Figure S3**). No other region in the genome exhibited as significant differentiation.

A haplotype block spanning ~1.75 Mb, from B4 positions 80,034,952 to 81,762,9005, was detected with a frequency of 0.5 across all cases (curly coated) and a frequency of 0.125 across the Selkirk Rex controls (straight coated). Within the block, from positions 81,210,239 to 81,789,564, a smaller haplotype block spanning ~600 Kb with a frequency of 0.89 within the cases was identified. The haplotype was unique to the curly Selkirk Rex phenotype and was not shared with the controls (**Figure S4**). Within the haplotype block 26 keratin genes were found based on human homology to the chromosomal region. Epithelial keratin genes and positions are presented in **Table S2**.

Population structure of the Persian-family breeds. A population structure analysis was conducted to examine the suitability of using sister breeds, Persian and Scottish Fold, as controls for the curly Selkirk Rex (**Figure 1**). To determine whether population stratification within the Persian-family breeds (Selkirk Rex, Persian, and Scottish Fold) was significant and would affect the GWAS, a known genetically distant breed (Birman) was assessed for population

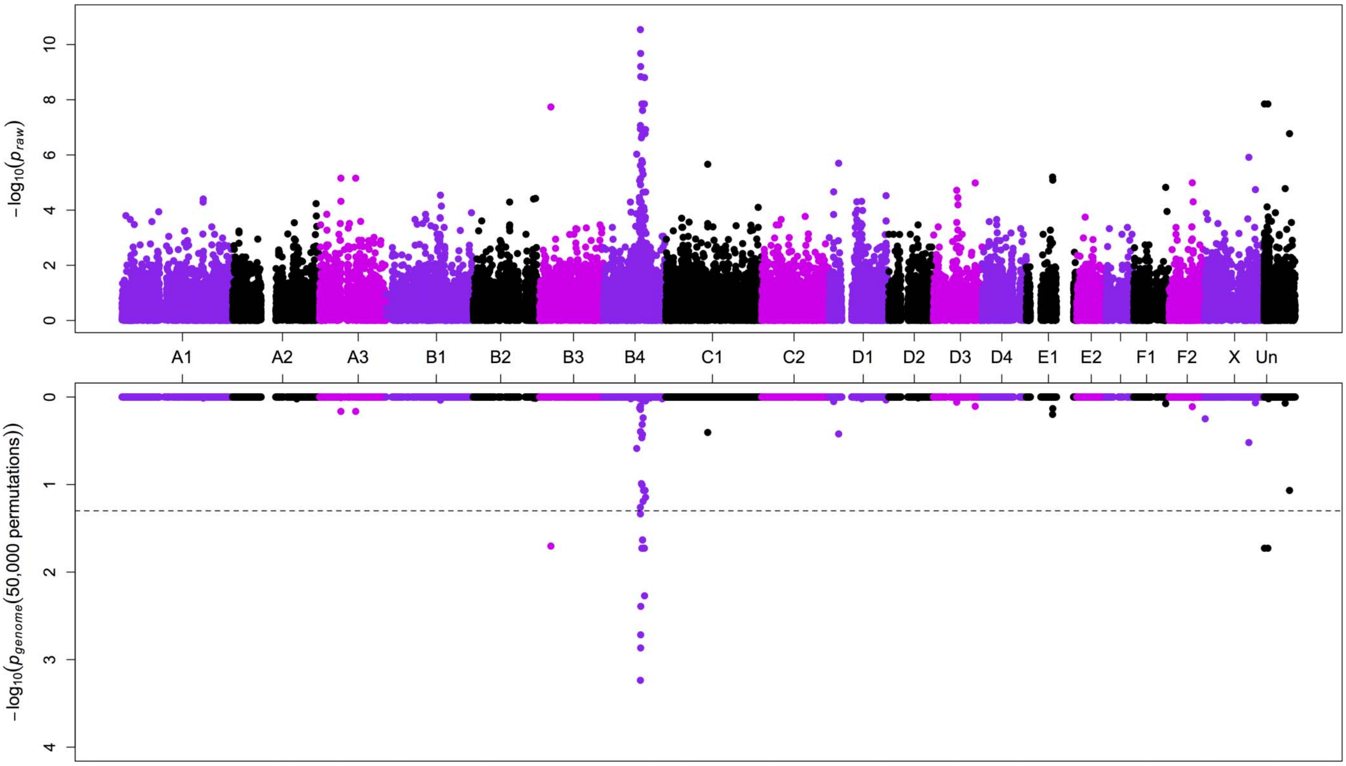

Figure 2 | Manhattan plot of the Selkirk Rex GWAS. The plot represents the P_{raw} (top) and P_{genome} (bottom) values of each SNP included in the case – control association study. The association study compared the curly coated Selkirk Rex with straight coated Selkirk Rex, Persian and Scottish Fold after exclusion of outliers (See Fig. S1) and highly related samples ($\widehat{P} > 0.3$), (See Fig. S2). A black dashed line suggests all the SNPs with genome-wide significance after permutations. A significant association with chromosome B4 was detected. The SNP with the highest association (B4.96694363) is at position 81,264,280 with a P_{raw} value of 2.87×10^{-11} and P_{genome} value of 0.00058.

structure with the aforementioned populations in a combined pool. The Birman breed was used as a phylogenetic control to test the hypothesis that the Persian-family of breeds are genetically indistinguishable in relation to a newly diverged lineage, such as the Selkirk Rex. As expected, when population structure was examined for $K = 2$ to $K = 8$, the best fit of the data across this set of breeds was represented by two groups ($K = 2$), the Birmans and the three Persian-family breeds (**Figure 3**)[27]. In contrast, differentiation between the Persian-family breeds at $K > 2$ was not well supported.

Candidate gene analyses (*KRT72, KRT73, KRT74*). A partial *KRT72* (GenBank accession no. KC178684) coding sequence (CDS), excluding exon 1, was analyzed in six unrelated curly Selkirk Rex, two American Wirehair, one LaPerm, and one control domestic shorthair. Ten variants were identified (**Table S3**), and none were responsible for an amino acid change. The intronic polymorphisms were highly variable and none of the polymorphisms segregated concordantly with the phenotype (data not shown). The entire *KRT73* coding sequence (GenBank accession no. KC189046) was obtained in three unrelated curly haired Selkirk Rex, two American Wirehair, one LaPerm and two domestic shorthairs. None of the variants segregated concordantly with the phenotype (**Table S4**). The entire *KRT74* genomic coding sequence and mRNA sequence was obtained (GenBank accession no. KC189047) in nine cats (four Selkirk Rex, two American Wirehair, two LaPerm, and one random bred cat). The mRNA was obtained from the hair bulb of two curly haired Selkirk Rex cats (homozygous and heterozygous) and a straight haired random bred cat. No identified polymorphisms within the gene were concordant with the phenotype and the mRNA sequence did not reveal any splice site variants. Moreover, none of the identified polymorphisms within the coding sequence of *KRT74* were associated with any other undiscovered rexoid phenotype within all the breeds included in sequenced cats (**Table S5**).

Table 1 | Ten highest associated chromosome B4 SNPs for Selkirk Rex curly hair GWAS

Chr.SNP*	Position	p-value
B4.96694363	81,264,280	2.87×10^{-11}
B4.97364183	81,789,564	2.10×10^{-10}
B4.97397960	81,813,919	2.23×10^{-10}
B4.97625142	81,980,475	1.45×10^{-9}
B4.107634573	90,684,713	1.58×10^{-8}
B4.100534818	84,831,461	1.42×10^{-8}
B4.101399061	85,512,084	1.42×10^{-8}
B4.102236759	86,189,301	1.42×10^{-8}
B4.103185944	86,988,687	1.42×10^{-8}
B4.103454785	87,195,353	1.42×10^{-8}

*The Chr.SNP location is based on cat assembly version 6.2/felCat5.

***KRT71* genomic and RNA analysis.** The entire CDS and mRNA sequence including partial 5′ UTR and 3′ UTR was obtained in two Selkirk Rex curly coated cats, one Selkirk Rex with straight hair and a domestic shorthair cat (**Table 2**). Genomic and RNA sequence was acquired as previously described[19]. A missense substitution at position 328 of the CDS causes an amino acid substitution, from a glutamic acid to a lysine. The mutation occurs in only one of the curly Selkirk Rex included in the sequencing. A second mutation detected genomically that causes a switch from a guanine to a cytosine at position c.445-1 of the CDS, likely disrupts the highly conserved acceptor splicing site of intron one. The mutation was present in

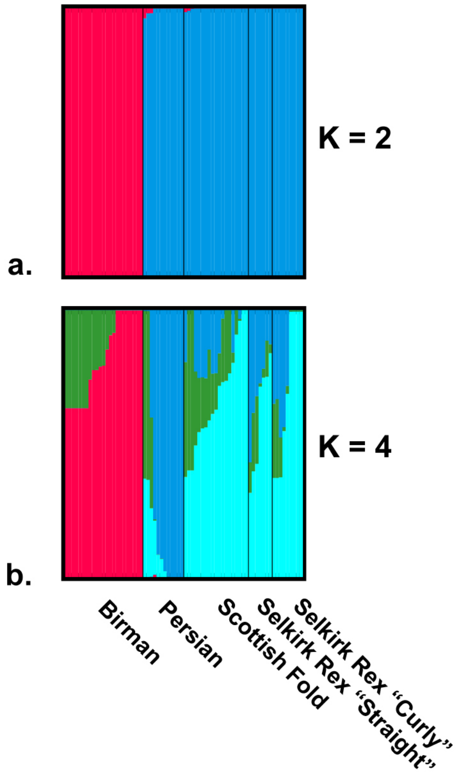

Figure 3 | Population clustering of the Persian-family cat breeds. *Admixture* population structure output of three Persian family breeds (Persian, Scottish Fold, and Selkirk Rex) and an out group represented by the Birman breed. (a) Significant population structure at $K = 2$ where a clear separate the control Birman (Eastern breed) from the Persian family breeds. (b) $K = 4$, where lack of clear differentiation between the Persian family breeds is observed.

the two curly coated Selkirk Rex and was absent in the straight hair Selkirk Rex, therefore was further investigated in detail. Sequence of the complete RNA transcript revealed that an alternative downstream acceptor was employed. The new alternative acceptor site was found within the first 18 bp of exon 2 (r.445_464del), thus an alternative splice site was recognized at base 18 in the modified exon 2. *In silico* translation predicted the loss of 6 amino acids in the *KRT71* protein (**Figure S5, Figure S6**).

***KRT71* genotyping analyses.** Fifty-nine unrelated straight and curly coated Selkirk Rex were genotyped for the identified c.445-1G > C polymorphism. All the curly coated Selkirk Rex (n = 47) were either homozygous or heterozygous for the identified mutation and the mutation was absent within the straight coated Selkirk Rex cats (n = 12). In addition, the genotyping was extended to several rexoid breeds, including American Wirehair, Cornish Rex, and Devon Rex, hairless breeds, including Sphynx, Peterbald, Kohana, straight coated breeds, including Persian and Scottish Fold, and random bred cats (**Table 3**). All the individuals (n = 108) were homozygous wildtype.

Discussion

The domestic cats' popularity and large population size permits the recognition of *de novo* phenotypes within the random bred population. In cats, new breed establishment is often due to artificial selection on a specific single gene trait, often responsible for an aesthetic pleasing phenotype, such as curly hair coats. These traits can reach fixation in the breed within only a few generations due to intense positive selection. This recent establishment of cat breeds and the strong selection of traits allows for powerful and highly efficient mapping of loci. One of the most recently developed breeds, the Selkirk Rex, is defined by an autosomal dominant rexoid coat mutation that occurred within the past ~10 generations[21]. The phenotype has complete penetrance and a clearly defined presentation. This curly coat phenotype was the original characteristic of the Selkirk Rex that defined the breed, and is not present or segregating in any other known cat breeds or populations. However, the LaPerm breed also has an autosomal dominant with a similar curly coated presentation.

In this study, to localize the rexoid phenotype of the Selkirk Rex, nine cases and 29 controls detected an association to cat chromosome B4. The recent origins of the breed and the positive selection for the trait in the cases produced extended linkage disequilibrium and haplotype blocks, making genome-wide association mapping amenable with fewer markers and fewer individuals. However, proper selection of controls was challenging because few Selkirk Rex are straight-coated since the phenotype is near fixation in the breed. In addition, Selkirk Rex are not a popular breed, with only 176 cats were registered in 2012 by the Cat Fanciers' Association (CFA), ranking 23rd of 42 breeds in popularity. In 2012, the Selkirk Rex breed had 68 registered litters, compared to the 4959 litters registered as the Persian breed, the most popular cat breed in the USA (www.cfa.org). The low number of individuals made ascertainment challenging. Straight-haired Selkirk Rex are rarely presented in cat shows, sometimes competing in the household pet category, hence ascertain was usually by solicitation of the breeder and not collection at cat shows. Therefore, because of an insufficient number of controls, straight-haired cats from different breeds were considered for the GWAS.

Several analyses were conducted to validate the controls for the Selkirk Rex GWAS. Previous population structure analysis placed the Selkirk Rex in a class of breeds termed the "Persian-family"[21]. This population cluster consists of five breeds: British shorthair, Exotic shorthair, Persian, Scottish Fold, and Selkirk Rex[21]. Persian and Scottish Fold members of the same breed family were included in the analyses as controls. These cats were fortuitously available from other array studies that required Persians and Scottish Folds. The admixture analysis supported the use of sister breeds as controls, and further suggested that little genetic stratification is evident in the individuals examined. Genome-wide F_{ST} analysis further demonstrated that the only significant difference between the Selkirk Rex curly group and the other control breeds is within the region that harbors the phenotypic trait under positive selection. Additional evidence is provided by the genomic inflation value, where the population sub-structure in the analyses is due to the inherent relatedness of the cats, rather than a different genetic background across samples.

Table 2 | *KRT71* variants in Selkirk Rex cats

Breed	Coat	c.78 T > G	c.189 T > G	c.328 G > A	c.384 A > C	c.445-1 G > C	c.445_464del	c.577 T > C
Selkirk Rex	Curly	T	C	R	C	S	Ins/del	C
Selkirk Rex	Curly	T	C	G	C	C	del/del	C
Selkirk Rex	Straight	T	C	G	C	G	Ins/Ins	C
Control	Straight	G	T	G	A	G	Ins/Ins	T
Amino Acid change		-	-	E > K	-	-	Δ8 AA	-

Table 3 | Genotypes of *KRT71* c.445-1G > C SNP in different cat breeds

Breed	Phenotype	No.	Genotype		
			C/C	C/G	G/G
American Wirehair	wirehair	4	0	0	4
Cornish Rex	curly	12	0	0	12
Devon Rex	curly	8	0	0	8
Kohana	hairless	5	0	0	5
LaPerm	curly	3	0	0	1
Persian	straight	16	0	0	16
Peterbald	hairless	7	0	0	7
Random Bred	straight	27	0	0	27
Scottish Fold	straight	12	0	0	12
Selkirk Rex	curly	47	11	36	0
	straight	12	0	0	12
Sphynx	hairless	9	0	0	9
Tennessee Rex	curly	2	0	0	2
Ural Rex	curly	3	0	0	3
Total		155	11	36	106

The significant association detected on cat chromosome B4 contained a cluster of keratin genes. Keratins consist of two groups, type I (acidic) and type II (neutral to basic), of intermediate filament proteins that can be found on human chromosomes 17 and 12 and on cat chromosomes E1 and B4, respectively[28–33]. Hair keratins form a strong unmineralized tissue within the HF controlling the hair structure and morphogenesis. Type I and II keratins form keratin intermediate filaments in the cytoplasm through heterodimerization and contribute to cytoskeleton formation[34,35]. The epithelial keratin family contains hair follicle-specific keratins (*KRT25, KRT26, KRT27, KRT28, KRT71, KRT72, KRT73, KRT74, KRT75*). The haplotype analysis revealed a unique haplotype, spanning ~600 Kb, exclusive to the curly coated Selkirk Rex. Of the extended keratin gene cluster, only the *KRT71, KRT72, KRT73, KRT74* and *KRT75* genes were found within the region of the unique haplotype. Due to their function and localization, the genes were considered excellent candidates for the Selkirk Rex phenotype.

A previous study[19], conducted on *KRT71* on rexoid cat breeds, including Selkirk Rex, discouraged an initial analysis of this gene and therefore *KRT71* was re-examined only after the analyses of *KRT74, KRT73* and *KRT72* was proving unsuccessful. Furthermore, several recent human studies implicated mutations within *KRT74* to be associated with autosomal dominant woolly hair[14,36]. Hence, *KRT74* was considered the highest priority candidate gene. After the characterizations of the *KRT74, KRT73* and *KRT72* keratins in the curly Selkirk Rex, the owners of the cats analyzed in the previous study were contacted to confirm phenotypes. The Selkirk Rex used in the previous study was ascertained at a cat show and entered in the laboratory database as a curly-coated Selkirk Rex since only curly coated Selkirk Rex can be shown for competition. The Selkirk Rex cat used in 2010 study actually competed in the household pet category and had straight hair, hence incorrectly phenotyped. Therefore, *KRT71* was incorrectly excluded as a candidate gene for the hair condition.

The identified *KRT71* splice variant within the Selkirk Rex breed is confirmed by genomic analysis and segregates concordantly with the phenotype in all 47 unrelated curly Selkirk Rex analyzed. *KRT71* CDS in cats is 1575 bp, which translates into a 524 amino acid protein. The keratin protein consists of three domains: an N-terminal head domain, a low-complexity coiled-coil C-terminal domain and the helix-forming α-helical rod domain. The N terminus and the C terminus of the α-helical rod domain are known as the helix initiation motif (HIM) and helix termination motif (HTM), domains extremely conserved and critical in forming heteropolymers of specific type I and type II cytokeratin[34,37]. Most pathogenic mutations for keratin diseases have been identified within either the HIM of the HTM domains of various keratins. In mice, several mutations within the HIM domain (Ca^{J}, Ca^{9J}, Ca^{10J}, Ca^{Rin}) and the HTM domain (Ca^{medJ}, *Cal4*) are responsible for the *Caracul* (*Ca*) and the *Caracul-like 4* (*Cal4*) dominant curly phenotypes[7,38,39]. In a recent study, a missense mutation detected within the HIM domain of the *KRT71* protein underlies autosomal dominant woolly hair and hypotrichosis in human[13]. Other dominant mutations in close proximity to the HIM and HTM domains are also reported in dogs and rats[11,40]. As suggested by the human study[13], the lack of the six amino acids within the α-helical rod domain of the cat, from positions 149 to 154 of the *KRT71* protein, may disrupt the endogenous keratin intermediate filament formation in a dominant-negative manner. The identified cat *KRT71* variant is unique to Selkirk Rex and was not identified in any other rexoid breed. Although the polymorphism is segregating concordantly with the phenotype in all the cats in this study and has not been identified in any other curly coated or straight-coated cats, the identified mutation may only represent a concordant variant in LD with the causative variant that would be elsewhere in the region.

Consistent with nomenclature of the mouse, *SADRE* was the suggested locus designation for the Selkirk autosomal dominant rex[21] hence $KRT71^{SADRE}$ is the suggested allele designation forming the allelic progression $KRT71^{SADRE} > KRT71^{+}$. Downstream mutations within the same domain are responsible for the autosomal recessive curly hair in Devon Rex, $KRT71^{re}$, and the naked phenotype of the Sphynx breed, $KRT71^{hr}$[19]. The Devon Rex coat lacks the guard hair and shows a reduction in length and thickness of the remaining fibers[41] while the Sphynx coat show a guard and thinner hairs that break easily conferring the naked phenotype[20]. Previous studies show that *KRT71* is dominantly expressed in the inner root sheath (IRS)[42] where specific type II keratins form heterodimers with IRS specific type I keratins[42–44]. Potentially, mutations in Devon Rex and Sphynx within the α-helical rod domain are recessive because the truncated *KRT71* protein can still form heterodimers incorporated in the hair shaft.

Three mutations in KRT71 have now been identified in cats, forming the allelic series, $KRT71^{SADRE} > KRT71^{+} > KRT71^{re} > KRT71^{hr}$. For the first time in cats, a mutation within *KRT71* is linked to an autosomal dominant form of curly hair coat. The identified variant should contribute to future studies elucidating the role of *KRT71* domains in hair follicle morphogenesis, development and hair growth in mammals.

Methods

Sampling. Private owners and breeders of Selkirk Rex cats voluntarily participated by donating DNA of their cats using buccal swabs. Sample collection was approved under the University of California - Davis, Institutional Animal Care and Use Committee (IACUC) animal protocols 16691 and 15933. One hundred cats with rexoid hair and or hairless phenotypes including, American wirehair, Cornish Rex, Devon Rex, Kohana, LaPerm, Peterbald, Selkirk Rex, Sphynx, Tennessee Rex and Ural Rex were ascertained as well as 67 cats with straight hairs, including Selkirk Rex, Scottish Fold, Persian and random bred cats). Additional straight coated cats were available from archived samples. Hairs from three Selkirk Rex cats and one random bred cat were plucked from random locations over the dorsum and scapular region, the hair bulbs were stored in RNAlater (Qiagen, Valencia, CA).

SNP array genotyping. Seventeen Selkirk Rex cases and 62 controls, which included seven Selkirk Rex straight haired controls, 23 Persian and 32 Scottish fold, were included in the GWAS. Genomic DNA was prepared as previously described[45]. Approximately 600 ng of genomic DNA was submitted to Neogene, Inc. (Lincoln, NE, USA) for genotyping on the illumina Infinium Feline 63 K iSelect DNA array (illumina, Inc., San Diego, CA).

Genome wide association study. SNP genotyping rate and minor allele frequency was evaluated using PLINK[46]. SNPs with a MAF < 5%, genotyping rate < 80%, and individuals genotyped for < 80% of SNPs were excluded from downstream analyses. An MDS with 2 dimensions was performed on 52,293 SNPs in PLINK to evaluate population substructure within cases and controls. Inflation of p-values was evaluated by calculating the λ and assessed with a Q-Q plot. The $\hat{P}$ for each individual was calculated using PLINK. The cats not tightly clustered in the MDS plot and the cats with $\widehat{P} > 0.3$ were removed from the analysis to reduce λ. A case-control association analysis was performed and corrected with 50,000 t-max permutations (-mperm 50,000). T-max permuted p-values were considered genome-wide significant at $p < 0.05$. A Manhattan plot of the results was generated using R (www.r-project.org/). Haplotypes were determined by considering SNPs from positions 79,949,675 to 85,062,355 of cat chromosome B4 and the identified block is presented as plots produced by HAPLOVIEW[47].

Population structure analysis. Bayesian clustering across populations (Birman, Persian, Scottish Fold, straight coated Selkirk Rex and curly coated Selkirk Rex) was evaluated with *Admixture*[48]. Populations were tested from K = 2 up to K = 8 (unsupervised), with 30 replicate runs. A separate random seed was generated for each K of each run, and cross-validation error was recorded for each result. The dataset consisted of "Persian Family" breeds represented by Selkirk Rex (both curly and straight), Scottish Fold, and Persian. Birman breed was used as an Eastern breed out-group. Sample sizes were 7, 9, 19, 12, and 23, respectively, with 43,488 autosomal SNPs included in the analysis. Thirty replicates of each Q file for each K were merged in CLUMPP[49]. The LargeKGreedy METHOD was used with GREEDY_OPTION 2 set to 2 (for random input orders). The merged Q files were then visualized in DISTRUCT to output the results for each K[50].

Genome-wide F_{ST} scans. The genome-wide analysis of F_{ST} values was performed using the *Pegas* package in R[51] for each SNP marker, comparing the curly Selkirk Rex population to Persian and Scottish Fold with all the other populations, independently and combined. Each SNP F_{ST} value was plotted as a function of its position along each autosomal chromosome within the genome.

***KRT72, KRT73, KRT74* genomic analyses.** The genomic analyses of *KRT72, KRT73, and KRT74* were conducted on genomic DNA from eight - nine cats including three – four cases and five controls, including curly coated cats of other breeds (**Table S3–S5**). Partial CDS of *KRT72, KRT73* and *KRT74 are* publicly available (http://ensembl.org) and can be found on GeneScaffold_3221:782,847-797,518:-1, GeneScaffold_3221:808,058-818,138:-1 and GeneScaffold_3221:764,347-771,930:-1, respectively. Alignment of the individual keratin genes' exon to the *Felis catus* – WGS trace archive database allowed for the retrieval of missing exonic sequence (http://blast.ncbi.nlm.nih.gov). Primers were designed in both UTRs and intronic regions, flanking the exons. Primers were tested for efficient product amplification on a DNA Engine Gradient Cycler (MJ Research, GMI, Ramsey, MN) and the final PCR magnesium concentrations, annealing temperatures, and amplicon sizes for each primer pair are shown in **Table S6**. PCR and thermocycling conditions were conducted as previously described[52]. The PCR products were purified with ExoSap (USB, Cleveland, OH) per the manufacturer's recommendations and directly sequenced using the BigDye terminator Sequencing Kit v3.1 (Applied Biosystems, Foster City, CA). Sequences were verified and aligned using the software *Sequencer* version 4.10 (Gene Codes Corp., Ann Arbor, MI).

KRT71 genomic and mRNA analysis. The genomic analysis of *KRT71* was conducted on genomic DNA from four cats including two curly Selkirk Rex, one straight hair Selkirk Rex and one control random bred cat (**Table 2**). The hair samples (same cats used for the genomic analysis) stored in RNAlater (Qiagen) were used to isolate RNA using PureLink RNA mini Kit (Invitrogen, Carlsbad, CA). The complete CDS of *KRT71* is publicly available (http://ensembl.org) and can be found on GeneScaffold_3221:736,547-745,383:-1. Primers were designed in the UTR regions, flanking the exons containing the full CDS and tested for efficient product amplification on a DNA Engine Gradient Cycler (MJ Research). Primer sequences and amplicon size for each primer pair are shown in **Table S6**. PCR and thermocycling conditions were conducted as previously described at 1.5 mM Mg^{2+} at 62°C[52]. The PCR products were purified, sequenced and aligned as previously described for the other keratins genomic analyses. Complementary DNA templates were synthesized by reverse transcription of 1 μg of mRNA from hair bulbs of a straight Selkirk Rex cat and two curly Selkirk Rex with the primers for the cDNA analysis provided in **Table S6**.

KRT71 mutation genotyping. To further correlate the identified mutation with the autosomal dominant rexoid hair, all Selkirk Rex cats from the array cohort (n = 17), Selkirk Rex within the laboratory archive (n = 35) and all the array straight haired Selkirk Rex (n = 7) were directly sequenced for the region containing the identified mutation. Additionally, Persians (n = 16) and Scottish Folds (n = 12) used in the array analysis were also genotyped for the identified mutation. Several curly coated and hairless breeds such as American Wirehair, Cornish Rex, Devon Rex, Kohana, LaPerm, Peterbald, Sphynx, Tennessee Rex and Ural Rex and 27 random bred cats were also included in the genotyping (**Table 3**). Primers for the direct sequencing genotyping of the implicated mutation are presented in **Table S6**.

1. Hardy, M. H. The secret life of the hair follicle. *Trends Genet* **8**, 55–61 (1992).
2. Millar, S. E. Molecular mechanisms regulating hair follicle development. *J Invest Dermatol* **118**, 216–225 (2002).
3. Paus, R. & Foitzik, K. In search of the "hair cycle clock": a guided tour. *Differentiation* **72**, 489–511 (2004).
4. Langbein, L. & Schweizer, J. Keratins of the human hair follicle. *Int Rev Cytol* **243**, 1–78 (2005).
5. Schmidt-Ullrich, R. & Paus, R. Molecular principles of hair follicle induction and morphogenesis. *Bioessays* **27**, 247–261 (2005).
6. Searle, A. G. & Jude, A. C. The rex type of coat in the domestic cat. *J. genet.* **54**, 506–512 (1956).
7. Kikkawa, Y. *et al.* A small deletion hotspot in the type II keratin gene *mK6irs1/Krt2-6g* on mouse chromosome 15, a candidate for causing the wavy hair of the *caracul* (*Ca*) mutation. *Genetics* **165**, 721–733 (2003).
8. Sundberg, J. P. *et al.* Angora mouse mutation: altered hair cycle, follicular dystrophy, phenotypic maintenance of skin grafts, and changes in keratin expression. *Vet Pathol* **34**, 171–179 (1997).
9. Ng, C. S. *et al.* The chicken frizzle feather is due to an alpha-keratin (*KRT75*) mutation that causes a defective rachis. *PLoS Genet* **8**(7), e1002748 (2012).
10. Kuramoto, T., Hirano, R., Kuwamura, M. & Serikawa, T. Identification of the rat Rex mutation as a 7-bp deletion at splicing acceptor site of the *Krt71* gene. *J Vet Med Sci* **72**, 909–912 (2010).
11. Cadieu, E. *et al.* Coat variation in the domestic dog is governed by variants in three genes. *Science* **326**, 150–153 (2009).
12. Sarkar, R., Koranne, R. V., Kakar, N. & Sardana, K. Hereditary wooly hair in an Indian family. *J Dermatol* **27**, 220–221 (2000).
13. Tariq, M., Azhar, A., Baig, S. M., Dahl, N. & Klar, J. A novel mutation in the *Lipase H* gene underlies autosomal recessive hypotrichosis and woolly hair. *Scientific Reports* **2**, 730 (2012).
14. Shimomura, Y., Wajid, M., Petukhova, L., Kurban, M. & Christiano, A. M. Autosomal-dominant woolly hair resulting from disruption of keratin 74 (KRT74), a potential determinant of human hair texture. *Am J Hum Genet* **86**, 632–638 (2010).
15. Azeem, Z. *et al.* Novel mutations in G protein-coupled receptor gene (P2RY5) in families with autosomal recessive hypotrichosis (LAH3). *Hum Genet* **123**, 515–519 (2008).
16. Robinson, R. The rex mutants of the domestic cat. *Genetica* **42**, 466–468 (1971).
17. Robinson, R. Oregon rex--a fourth rexoid coat mutant in the cat. *Genetica* **43**, 236–238 (1972).
18. Robinson, R. German rex: a rexoid coat mutant in the cat. *Genetica* **39**, 351–352 (1968).
19. Gandolfi, B. *et al.* The naked truth: Sphynx and Devon Rex cat breed mutations in KRT71. *Mamm Genome* **21**, 509–515 (2010).
20. Robinson, R. The Canadian hairless of Sphinx cat. *J Hered* **64**, 47–49 (1973).
21. Filler, S. *et al.* Selkirk Rex: morphological and genetic characterization of a new cat breed. *J Hered* **103**, 727–733 (2012).
22. Levy-Nissenbaum, E. *et al.* Hypotrichosis simplex of the scalp is associated with nonsense mutations in *CDSN* encoding corneodesmosin. *Nat Genet* **34**, 151–153 (2003).
23. Wen, Y. *et al.* Loss-of-function mutations of an inhibitory upstream ORF in the human hairless transcript cause Marie Unna hereditary hypotrichosis. *Nat Genet* **41**, 228–233 (2009).
24. Duzenli, S. *et al.* Identification of a *U2HR* gene mutation in Turkish families with Marie Unna hereditary hypotrichosis. *Clin Exp Dermatol* **34**, e953–956 (2009).
25. Shimomura, Y. *et al. APCDD1* is a novel Wnt inhibitor mutated in hereditary hypotrichosis simplex. *Nature* **464**, 1043–1047 (2010).
26. Fujimoto, A. *et al.* A missense mutation within the helix initiation motif of the *keratin K71* gene underlies autosomal dominant woolly hair/hypotrichosis. *J Invest Dermatol* **132**, 2342–2349 (2012).
27. Alexander, D. H. & Lange, K. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC bioinformatics* **12**, 246 (2011).
28. Lessin, S. R., Huebner, K., Isobe, M., Croce, C. M. & Steinert, P. M. Chromosomal mapping of human keratin genes: evidence of non-linkage. *J Invest Dermatol* **91**, 572–578 (1988).
29. Romano, V. *et al.* Chromosomal assignments of human type I and type II cytokeratin genes to different chromosomes. *Cytogenet Cell Genet* **48**, 148–151 (1988).
30. Rosenberg, M., RayChaudhury, A., Shows, T. B., Le Beau, M. M. & Fuchs, E. A group of type I keratin genes on human chromosome 17: characterization and expression. *Mol Cell Biol* **8**, 722–736 (1988).
31. Waseem, A., Alexander, C. M., Steel, J. B. & Lane, E. B. Embryonic simple epithelial keratins 8 and 18: chromosomal location emphasizes difference from other keratin pairs. *New Biol* **2**, 464–478 (1990).
32. Rogers, M. A. *et al.* Sequence data and chromosomal localization of human type I and type II hair keratin genes. *Exp Cell Res* **220**, 357–362 (1995).
33. Rogers, M. A. *et al.* Characterization of new members of the human type II keratin gene family and a general evaluation of the keratin gene domain on chromosome 12q13.13. *J Invest Dermatol* **124**, 536–544 (2005).
34. Coulombe, P. A. & Omary, M. B. 'Hard' and 'soft' principles defining the structure, function and regulation of keratin intermediate filaments. *Curr Opin Cell Biol* **14**, 110–122 (2002).
35. Moll, R., Divo, M. & Langbein, L. The human keratins: biology and pathology. *Histochemistry and cell biology* **129**, 705–733 (2008).
36. Wasif, N. *et al.* Novel mutations in the *keratin-74* (*KRT74*) gene underlie autosomal dominant woolly hair/hypotrichosis in Pakistani families. *Hum Genet* **129**, 419–424 (2011).

37. Hatzfeld, M. & Weber, K. The coiled coil of *in vitro* assembled keratin filaments is a heterodimer of type I and II keratins: use of site-specific mutagenesis and recombinant protein expression. *J Cell Biol* **110**, 1199–1210 (1990).
38. Peters, T. *et al*. Alopecia in a novel mouse model RCO3 is caused by mK6irs1 deficiency. *J Invest Dermatol* **121**, 674–680 (2003).
39. Runkel, F. *et al*. Morphologic and molecular characterization of two novel Krt71 (*Krt2-6g*) mutations: *Krt71rco12* and *Krt71rco13*. *Mamm Genome* **17**, 1172–1182 (2006).
40. Kuramoto, T., Hirano, R., Kuwamura, M. & Serikawa, T. Identification of the rat rex mutation as a 7-bp deletion at splicing acceptor site of the *Krt71* Gene. *J Vet Med Sci* **72**(7), 909–12 (2010).
41. Robinson, R. Devon Rex - a Third Rexoid Coat Mutant in Cat. *Genetica* **40**, 597–499 (1969).
42. Langbein, L., Rogers, M. A., Praetzel, S., Winter, H. & Schweizer, J. K6irs1, K6irs2, K6irs3, and K6irs4 represent the inner-root-sheath-specific type II epithelial keratins of the human hair follicle. *J Invest Dermatol* **120**, 512–522 (2003).
43. Langbein, L. *et al*. *K25* (*K25irs1*), *K26* (*K25irs2*), *K27* (*K25irs3*), and *K28* (*K25irs4*) represent the type I inner root sheath keratins of the human hair follicle. *J Invest Dermatol* **126**, 2377–2386 (2006).
44. Tanaka, S. *et al*. Mutations in the helix termination motif of mouse type I IRS keratin genes impair the assembly of keratin intermediate filament. *Genomics* **90**, 703–711 (2007).
45. Gandolfi, B. *et al*. First *WNK4*-hypokalemia animal model identified by genome-wide association in Burmese cats. *PLoS One* **7**, e53173 (2012).
46. Purcell, S. *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559–575 (2007).
47. Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265 (2005).
48. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* **19**, 1655–1664 (2009).
49. Jakobsson, M. & Rosenberg, N. A. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801–1806 (2007).
50. Rosenberg, N. A. Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes* **4**, 137–138 (2004).
51. Paradis, E. Pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* **26**, 419–420 (2010).
52. Bighignoli, B. *et al*. *Cytidine monophospho-N-acetylneuraminic acid hydroxylase* (*CMAH*) mutations associated with the domestic cat AB blood group. *BMC Genet* **8**, 27 (2007).

Acknowledgments

This work was supported by funding from the National Center for Research Resources R24 RR016094 and is currently supported by the Office of Research Infrastructure Programs/ OD R24OD010928, the Winn Feline Foundation (W10-014 and W10-015), the Cat Health Network grant D12FE-508, and the George and Phyllis Miller Feline Health Fund, Center for Companion Animal Health, School of Veterinary Medicine, University of California, Davis (2008-36-F). The authors would like to thank Ms. Donna Bass and Alessandra Piersanti for providing access to their cats for photos, hair samples and discussions about the breed. We also appreciate the technical support of Robert A. Grahn, Grace Lan, Carlyn B. Petersen, and Alejandro Cortes.

Author contributions

B.G. and L.A.L. conceived of the idea and planned the experiments. B.G. wrote the paper and L.A.L. provided the experimental supplies and S.F. and G.B. contributed with samples. S.E.K.J. and B.G. performed the experiments and H.A., R.K. and B.G. supported data analysis. L.A.L. and H.A. assisted with manuscript preparation.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Gandolfi, B. *et al*. A splice variant in *KRT71* is associated with curly coat phenotype of Selkirk Rex cats. *Sci. Rep.* **3**, 2000; DOI:10.1038/srep02000 (2013).