

Rodent *Lce* Gene Clusters; New Nomenclature, Gene Organization, and Divergence of Human and Rodent Genes

Journal of Investigative Dermatology advance online publication, 5 April 2007; doi:10.1038/sj.jid.5700792

TO THE EDITOR

Human and rodent *LCE* gene clusters are located on the epidermal differentiation complexes (EDC; 1q21 in human, 3F2.1 in mouse and 2q34 in rat) and encode multiple small genes with similarities, particularly over their N-terminal region, to small proline-rich proteins. Many *LCE* genes contain the glycine-serine-cysteine-rich motif typical of cornified envelope proteins such as loricrin (Zhao and Elder, 1997; Marshall *et al.*, 2001; Wang *et al.*, 2001; Jackson *et al.*, 2005). Their N- and C-termini are similar to known transglutaminase substrates and *LCE* proteins are demonstrated cornified envelope constituents (Marshall *et al.*, 2001; Steinert *et al.*, 2003). *LCE* proteins upregulate in the loricrin null mouse, suggesting that they can functionally substitute for loricrin (Koch *et al.*, 2000; Hohl, 2005). We report here major dissimilarity between human and rodent *LCE* gene clusters, with a surprising level of divergence between mouse and rat.

The human 17 gene *LCE* cluster comprises three major groups in three chromosomal clusters. Group members encode similar proteins with related expression patterns – thus human group 1 and 2 genes express mainly in external epithelia such as skin, whereas group 3 genes express usually in internal stratum corneum-forming epithelia (e.g. tongue surface) (Jackson *et al.*, 2005). Recently, a nomenclature was agreed for the 17 human *LCE* genes (formerly *XP5*, *EIG*, *SPRL*, *SPRRL*, and *LEP* genes) (Jackson *et al.*, 2005).

We present here a related agreed nomenclature for mouse and rat *Lce* genes (Table 1 and Figure 1). All work reported here was institutionally approved.

Rodent and human *LCE* gene expansion appears to have been largely independent (i.e. after the ancestral species diverged), and/or there has been rodent and human-specific gene homogenization after ancestral species divergence, probably via gene conversion. Rodents have 20–22 *Lce* genes, which cluster into two chromosomal groups (Figure 1). On the basis of conserved motifs rodent genes can be assigned into two groups (groups 1 and 3), which map to the two chromosomal clusters (Figure 1 and Figure S1). Rodent groups 1 and 3 are orthologs of human groups 1 and 3 (Figure S1; Jackson *et al.*, 2005). However, individual rodent *Lce* genes cannot be assigned as direct orthologs of human *LCE* genes. For example, human *LCE1A* is not a direct ortholog of rodent *Lce1a* (or does not clearly derive from a recent common ancestor, see Figure S1).

Rodent group 1 genes are characterized by an additional, variable length glycine-cysteine-serine cluster in their C-terminal portion, which is absent in human group 1 genes (Figure S1). The rodent group 1 expansion appears to be so recent that there is incomplete conservation between mouse and rat. The new nomenclature incorporates this difficulty as (1) clear mouse-rat orthologs have the same name (e.g. mouse and rat *Lce1b* appear to be orthologs; (2) paralogs (same species

gene duplications) have an additional digit, for example, murine *Lce1a1* and *Lce1a2* probably derive from a gene duplication occurring after mouse and rat diverged (Figure 1 and Table 1); and (3) genes which appear structurally dissimilar have independent names.

Rodents lack human group 2 *LCEs* (Figure S1; Jackson *et al.*, 2005). The specific function of human group 2 proteins is unknown, although their expression in human keratinocytes is particularly calcium-sensitive and, like human group 1 proteins, they express in skin (Jackson *et al.*, 2005; Su *et al.*, 2004). Absence of group 2 rodent genes is accompanied by an expansion of group 1 genes relative to human (Table 1 and Figure 1). Human group 3 *LCE* genes and rodent group 3 *Lce* genes are structurally similar (Figure S1), cluster in a similar position on the genome in relationship to *CRCT1/Crct1* (alias *NICE-1/Nice1*; Marenholz *et al.*, 2001; Figure 1), and both groups are expressed strongly in internal epithelia (Su *et al.*, 2004).

We speculated that the major differences between human and rodent *LCE* genes may reflect the difference in stratum corneum and barrier requirements between hair-covered skin and sparsely haired skin. However, chimpanzee (a primate with hair-covered skin) contains *LCE* group 2 genes and their group 1 gene organization is similar to human (Chimpanzee Sequencing and Analysis Consortium, 2005).

The surprising divergence in *LCE* chromosomal organization and protein coding sequences between species is consistent with the recent description of this gene cluster as an example of a

Abbreviations: *LCE*, late-cornified envelope; *Sprr*, small proline-rich

Table 1. Rodent *Lce* gene nomenclature

(b) Murine *Lce* gene nomenclature and chromosomal position

	Position on chromosome-3 ¹		Aliases, previous nomenclature	RIKEN	Accession	Predicted Protein
Lce1a1	92,732,614	92,732,145	Sprrl3	2200008B06Rik	NM_025984	NP_080260
Lce1b	92,741,673	92,741,186	Sprrl5	1110029C13Rik	NM_026822	NP_081098
Lce1a2	92,754,702	92,754,251	Sprrl2	1110004E04Rik	NM_028625	NP_082901
Lce1c	92,765,671	92,766,114		1110014K05Rik	NM_028622	NP_082898
Lce1d	92,771,553	92,771,055	Sprrl7	2310037L11Rik	NM_027137	NP_081413
Lce1e	92,793,485	92,793,016		1110031B11Rik	NM_026811	NP_081087
Lce1f	92,804,795	92,804,338		1110055J05Rik	NM_026394	NP_080670
Lcep1-ps	92,823,248	92,823,688				
Lce1g	92,836,446	92,835,965		1110058A15Rik	NM_025413	NP_079689
Lce1h	92,849,290	92,848,812	Sprrl9	2310066F03Rik	NM_026335	NP_080611
Lce1i	92,863,314	92,862,863		2310069N01Rik	NM_029667	NP_083943
Lce1j	92,874,872	92,874,453		2210405D07 ²		
Lce1k	92,892,278	92,891,901		2210013J22 ²		
Kprp ³	92,911,173	92,908,477		1110001M24Rik	NM_028629	NP_082905
Lce1l	92,935,996	92,935,503		1110008K04Rik	NM_028628	NP_082904
Lce3a	93,011,194	93,010,898		CJ238734 ²	NM_001039594	NP_001034683
Lce3b	93,019,003	93,019,299	Sprrl6A	2310007F04Rik	NM_025501	NP_079777
Lce3c	93,030,637	93,030,933	Sprrl1, Eig3	2300007B01Rik	NM_033175	NP_149410
Lce3d	93,043,483	93,043,779				
Lce3e	93,053,142	93,053,438				
Lce3f	93,078,277	93,078,573		2310002A05Rik	NM_001018079	NP_001018089
Crct1 ⁴	93,100,253	93,099,610	Nice-1, C1orf42	2300002G24Rik	NM_028798	NP_083074
Lce1m	93,104,053	93,103,521	Sprrl10, Lce5a	1110059L13Rik	NM_025420	NP_079696

(b) Rat *Lce* gene nomenclature and chromosomal position

	Position on chromosome 2 ⁵		Aliases, previous nomenclature	Accession	Predicted protein
Lce1a	185642402	185641980			
Lce1b	185647918	185647505	LOC685972	XM_001066001	XP_001066001
Lce1d	185677116	185677522	LOC685981	XM_001066061	XP_001066061
Lce1c	185684635	185684186	LOC685998	XM_001066116	XP_001066116
Lce1n	185702807	185702386			
Lce1f	185715540	185715172	RGD1565626	XM_579999	XP_579999
Lce1o	185732316	185731927			
Lce1p	185754161	185753802			
Lcep1-ps	185763742	185764070			
Lce1q	185765954	185765505			
Lce1r	185777169	185776777			
Lce1s	185791018	185790629	LOC686064	XM_001066389	XP_001066389
Lce1t	185808428	185808057			
Lce1k	185827899	185827497			
Kprp ³	185843694	185839396	RGD1303244	NM_001002290	NP_001002290
Lce1l	185872621	185872196	LOC686125	XM_001066620	XP_001066620

Table 1. continued

(b) Rat Lce gene nomenclature and chromosomal position

	Position on chromosome 2 ⁵		Aliases, previous nomenclature	Accession	Predicted protein
Lce3a	185956444	185956175			
Lce3g	185964395	185964691			
Lce3h	185975427	185975723			
Lce3i	185992866	185993162			
Lce3j	186005759	186006055			
Lce3e	186013398	186013694			
Lce3f	186031610	186031906			
Crct1 ⁴	186049846	186049592	LOC688401, C1orf42, Nice-1	XM_001066782	XP_001066782
Lce1m	186053823	186053314	LOC688413	XM_001066832	XP_001066832

¹Chromosome position from NCBI Build 36.1, Contig NT 39240.6, numbers refer to translation start and stop positions.

²Riken full-length clones unattached to gene.

³Lee *et al.* (2005).

⁴Marenholz *et al.* (2001).

⁵Chromosome position from NCBI, RGSCv3.4, NW_047626.2, numbers refer to translation start and stop positions.

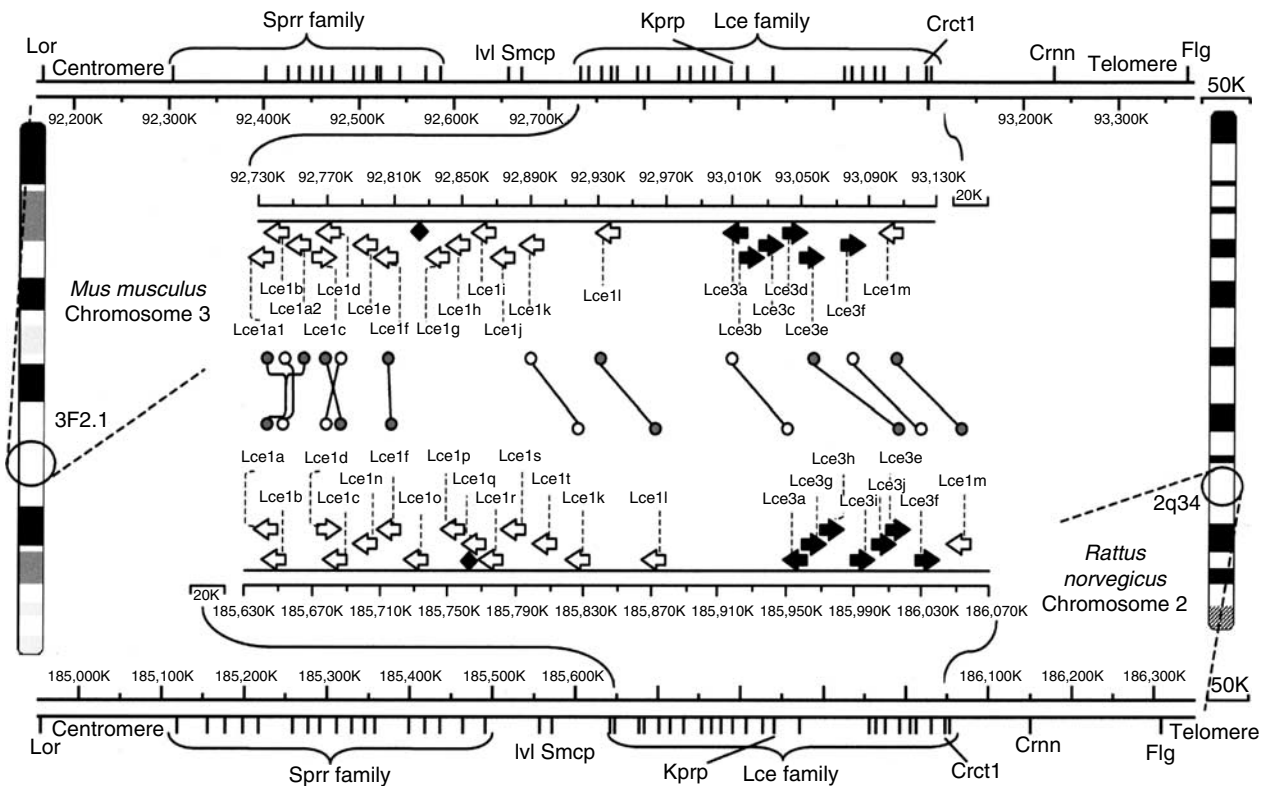


Figure 1. Chromosomal regions from the rodent epidermal differentiation complexes between lorincrin and filaggrin are shown. The magnified insets show the Lce gene clusters. Clear arrows indicate group 1 genes and black arrows group 3 genes, with the arrow denoting the direction of transcription. Pseudogenes are shown by dark grey diamonds. Homologous genes are indicated by the central graphic. Crct1, cysteine-rich C-terminal 1; Crnn, cornulin; Flg, filaggrin; Lvl, involucrin; Kgrp, keratinocyte expressed proline-rich protein; Smcp, sperm mitochondria-associated cysteine-rich protein; Lce, late-cornified envelope; Lor, lorincrin; Sprr, small proline-rich. Chromosomal positions for mouse are from NCBI Build 36.1, Contig NT 39240.6, whereas chromosomal positions for rat are from NCBI, RGSCv3.4, NW_047626.2.

“rapidly evolving” gene group (Varki and Altheide, 2005) or one of the gene groups that may define differences between species. The apparent dispensability of group two genes in rodents is consistent with the recent findings of LCE gene deletion heterogeneity. Within a small sample of Caucasian human patients, one patient was recently shown to lack expression of a human group one gene resulting from an LCE1D/E region deletion (Jackson et al., 2005; Tilli and Byrne, unpublished observation). Furthermore, LCE3C was recently shown to be a common deletion polymorphism in humans, with deletion frequency varying between racial groups (McCarroll et al., 2006).

The large number of LCE genes encoding very similar proteins and the apparent dispensability of specific genes within species and an entire gene group across species strongly suggests redundant protein function, invoked as a backup for stratum corneum function (Steinert, 2000). The advantage, if any, associated with LCE gene amplification in primates and rodents may reside in the coamplification of gene regulatory sequences that permit more flexible change to barrier characteristics in response to a range of environmental stimuli (Cabral et al., 2001; Jackson et al., 2005).

CONFLICT OF INTEREST

The authors state no conflict of interest.

ACKNOWLEDGMENTS

This research was supported by grants from St Bartholomew and The London Charitable Trust and Cancer Research UK (CB).

Stuart J. Brown¹, Claudia M.L.J Tilli^{1,2}, Ben Jackson¹, Ariel A. Avilion^{1,2}, Mike C. MacLeod³, Lois J. Maltais⁴, Ruth C. Lovering⁵ and Carolyn Byrne^{1,2}

¹Institute for Cell and Molecular Science, Barts and The London, Queen Mary School of Medicine and Dentistry, University of London, London, UK; ²Cancer Research UK Skin Tumour Lab, Barts and The London, Queen Mary School of Medicine and Dentistry, University of London, London, UK; ³Department of Carcinogenesis, University of Texas M. D. Anderson Cancer Center, Smithville, Texas, USA; ⁴Mouse Genomic Nomenclature Committee (MGNC), Mouse Genome Informatics (MGI) Resource, The Jackson Laboratory, Bar Harbor, Maine, USA; ⁵HUGO Gene Nomenclature Committee (HGNC), Department of Biology, University College London, London, UK. E-mail: C.R.Byrne@qmul.ac.uk

SUPPLEMENTARY MATERIAL

Materials and Methods.

Figure S1. Alignment of human (h), mouse (m) and rat (r) LCE predicted coding sequences which highlight the similarities and conserved motifs.

REFERENCES

- Cabral A, Voskamp P, Cleton-Jansen AM, South A, Nizetic D, Backendorf C (2001) Structural organization and regulation of the small proline-rich family of cornified envelope precursors suggest a role in adaptive barrier function. *J Biol Chem* 276:19231–7
- Chimpanzee Sequencing and Analysis Consortium (2005) Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* 437:69–87
- Hohl D (2005) Formation of the cornified envelope. *Exp Dermatol* 14:777–80
- Jackson B, Tilli CL, Hardman M, Avilion A, Macleod M, Ashcroft G et al. (2005) Late cornified envelope family in differentiating epithelia – response to calcium and ultraviolet irradiation. *J Invest Dermatol* 124:1062–70

- Koch PJ, De Viragh PA, Schärer E, Bundman D, Longley MA, Bickenbach J et al. (2000) Lessons from loricrin-deficient mice: compensatory mechanisms maintaining skin barrier function in the absence of a major cornified envelope protein. *J Cell Biol* 151:389–400
- Lee WH, Jang S, Lee JS, Lee Y, Seo EY, You KH et al. (2005) Molecular cloning and expression of human keratinocyte proline-rich protein (hKPRP), an epidermal marker isolated from calcium-induced differentiating keratinocytes. *J Invest Dermatol* 125:995–1000
- Marenholz I, Zirra M, Fischer DF, Backendorf C, Ziegler A, Mischke D (2001) Identification of human epidermal differentiation complex (EDC)-encoded genes by subtractive hybridization of entire YACs to a gridded keratinocyte cDNA library. *Genome Res* 11:341–55
- Marshall D, Hardman MJ, Nield KM, Byrne C (2001) Differentially expressed late constituents of the epidermal cornified envelope. *Proc Natl Acad Sci USA* 98:13031–6
- McCarroll SA, Hadnott TN, Perry GH, Sabeti PC, Zody MC, Barrett JC (2006) Common deletion polymorphisms in the human genome. *Nat Genet* 38:86–92
- Steinert PM (2000) The complexity and redundancy of epithelial barrier function. *J Cell Biol* 151:F5–8
- Steinert PM, Parry DA, Marekov LN (2003) Trichohyalin mechanically strengthens the hair follicle: multiple cross-bridging roles in the inner root sheath. *J Biol Chem* 278:41409–19
- Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D et al. (2004) A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc Natl Acad Sci USA* 101:6062–7
- Varki A, Altheide TK (2005) Comparing the human and chimpanzee genomes: searching for needles in a haystack. *Genome Res* 15:1746–58
- Wang A, Johnson DG, MacLeod MC (2001) Molecular cloning and characterization of a novel mouse epidermal differentiation gene and its promoter. *Genomics* 73:284–90
- Zhao XP, Elder JT (1997) Positional cloning of novel skin-specific genes from the human epidermal differentiation complex. *Genomics* 45:250–8