Genetic tagging of humpback whales

Palsboll, PJ; Allen, J; Berube, M; Clapham, PJ; Feddersen, TP; Hammond, PS; Hudson, RR; Jorgensen, H; Katona, S; Larsen, AH

Published in:
Nature

DOI:
10.1038/42005

Publication date:
1997

25. Stanley, H. H. *et al.* Statistical mechanics in biology: how ubiquitous are long-range correlations? *Physica A* **205,** 214–253 (1996).
26. Plotnick, R. E. & McKinney, M. Evidence of self-organization in planktic foraminiferal evolution: implications for interconnectedness of paleoecosystems. *Palaios* **8,** 202–212 (1993).
27. Burlando, B. The fractal geometry of evolution. *J. Theor. Biol.* **163,** 161–172 (1993).
28. Solé, R. V., Bascompte, J. & Manrubia, S. C. Extinctions: Bad genes or weak chaos? *Proc. R. Soc. Lond. B* **263,** 1407–1413 (1996).
29. Raup, D. M. & Sepkoski, J. J. Jr Mass extinctions and the marine fossil record. *Science* **215,** 1501–1503 (1982).
30. Glen, W. (ed.) *The Mass Extinction Debates* (Stanford Univ. Press, 1984).
31. Sepkoski, J. J. Jr A kinetic model of Phanerozoic taxonomic diversity III. Post-Paleozoic families and mass extinctions. *Paleobiology* **10,** 246–267 (1984).
32. Newman, M. E. J. *A Model of Mass Extinction* (Wkg Pap. 97-02-013 Santa Fe Institute, 1997); *J. Theor. Biol.* (in the press).
33. Signor, P. W. & Lipps, J. H. Sampling bias, gradual extinction patterns and catastrophes in the fossil record. *Geol. Soc. Am. Spec. Pap.* **190,** 291–296 (1982).

Acknowledgements. We thank J. Bascompte, M. Newman, J. Pérez-Mercader, C. Patterson, B. Goodwin, S. Kauffman, K. Sneppen. T. Keitt, J. Sepkoski and M. Gell-Mann for help at different stages of this work and for discussions. This work was supported by DGYCIT (R.V.S. and S.C.M.), CIRIT (S.C.M.), a Leverhulme grant (M.J.B.), the division of Materials Science (P.B.) and the Santa Fe Institute (R.V.S. and P.B.).

Correspondence and requests for materials should be addressed to R.V.S. (e-mail: sole@santafe.edu).

Genetic tagging of humpback whales

Per J. Palsbøll*, Judith Allen†, Martine Bérubé*‡, Phillip J. Clapham§, Tonnie P. Feddersen*, Philip S. Hammond‖, Richard R. Hudson¶, Hanne Jørgensen*, Steve Katona†, Anja Holm Larsen*, Finn Larsen#, Jon Lien☆, David K. Mattila§, Jóhann Sigurjónsson, Richard Sears††, Tim Smith‡‡, Renate Sponer*, Peter Stevick† & Nils Øien§§**

* *Department of Population Biology, Universitetsparken 15, Copenhagen DK-2100, Denmark*
† *Allied Whale, College of the Atlantic, Bar Harbor, Maine 04609, USA*
‡ *Department of Natural Resources Science, Ste-Anne de Bellevue, Quebec H9X 3V9, Canada*
§ *Center for Coastal Studies, 59 Commercial Street, Provincetown, Massachusetts 02657, USA*
‖ *Sea Mammal Research Unit, University of St Andrews, St Andrews, Fife KY16 8LB, UK*
¶ *Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92697, USA*
# *Greenland Institute of Natural Resources, PO Box 570, 3900 Nuuk, Greenland*
☆ *Whale Research Group, Memorial University, Newfoundland A1B 3X9, Canada*
** *Marine Research Institute, Skulagata 4, PO Box 1390, 121 Reykjavik, Iceland*
†† *Mingan Island Cetacean Study Inc., 285 Green Street, Ste-Lambert, Quebec J4P 1T3, Canada*
‡‡ *US National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, Massachusetts 02543, USA*
§§ *Institute of Marine Research, P.O. Box 1870, Nordnes, N1-15024 Bergen, Norway*

The ability to recognize individual animals has substantially increased our knowledge of the biology and behaviour of many taxa[1]. However, not all species lend themselves to this approach, either because of insufficient phenotypic variation or because tag attachment is not feasible. The use of genetic markers ('tags') represents a viable alternative to traditional methods of individual recognition, as they are permanent and exist in all individuals. We tested the use of genetic markers as the primary means of identifying individuals in a study of humpback whales in the North Atlantic Ocean. Analysis of six microsatellite loci[2,3] among 3,060 skin samples collected throughout this ocean allowed the unequivocal identification of individuals. Analysis of 692 'recaptures', identified by their genotype, revealed individual local and migratory movements of up to 10,000 km, limited exchange among summer feeding grounds, and mixing in winter breeding areas, and also allowed the first estimates of animal abundance based solely on genotypic data. Our study demonstrates that genetic tagging is not only feasible, but generates data (for example, on sex) that can be valuable when interpreting the results of tagging experiments.

Skin biopsy[4] or sloughed skin[5] samples from free-ranging humpback whales (*Megaptera novaeangliae*) were collected across the North Atlantic between 1988 and 1995. Total-cell DNA was extracted[6], and the sex[7] and genotype at six mendelian inherited[8] microsatellite loci[9] were determined for each sample. From the 3,060 samples analysed, we detected 2,368 unique genotypes. The expected number of samples collected from different individuals with identical genotype arising by chance was estimated at less than one (see Methods). Because of this, and the fact that all samples with identical genotypes were of consistent sex, we believe that the 3,060 samples represented 2,368 individual whales. Of the 692 recaptures observed during the study, 216 occurred on the summer feeding grounds (Fig. 1). Of these, 96% ($n = 207$) occurred within the same feeding area (Fig. 1), confirming previous behavioural[10,11] and genetic[12,13] observations of maternally directed fidelity to specific feeding grounds. The remaining 4% ($n = 9$) of the recaptures on the summer feeding grounds were detected in different but adjacent feeding grounds (Fig. 1). However, significantly more recaptures were detected within these sampling areas than would be expected if the areas constituted one intermixing feeding aggregation and is consistent with the notion of maternally directed site fidelity (see Methods).

Of the 114 individuals recorded on both summer feeding and winter breeding grounds (Fig. 1), two had migrated from the West Indies to either Jan Mayen or Bear Island (in the Barents Sea). These genetic recaptures represent the most extensive one-way movements (6,435 and 7,940 km, respectively) recorded in this study, and support recent findings[13,14] that whales feeding in the Barents Sea share a common breeding ground with other North Atlantic humpback whales. In three other individuals, which were each sampled on three occasions, movements were documented from a feeding ground to the breeding range and back, involving minimum migration distances of up to 10,000 km between the first and last sampling event. No feeding ground was disproportionately represented among the recaptures in the West Indies (Fig. 1), supporting the current view that humpback whales in the North Atlantic constitute a single panmictic population[13,15,16] (G-test, $G_{5\text{d.f.}} = 4.68$, $P < 0.46$).

As with traditional identification methods[1], microsatellite data lend themselves to abundance estimation using mark–recapture statistical methods[17], although to our knowledge this has not previously been attempted. Using breeding-ground samples collected during 1992 and 1993, we estimated the North Atlantic humpback whale population at 4,894 (95% confidence interval, 3,374–7,123) males and 2,804 (95% confidence interval, 1,776–4,463) females. This total of 7,698 whales is substantially (albeit not significantly) higher than the most recent previous photographic-based estimate of 5,505 (ref. 10) (95% confidence interval, 2,888–8,122). Preliminary results from new and more reliable photographic estimates are also larger than previous estimates (T.S. *et al.*, manuscript in preparation), and could partly be due to population growth during the intervening decade since the previous estimate[18]. The significantly different estimates for males and females are unexpected given the even sex ratio observed on the feeding grounds[19] (Table 1) and among 198 calves that we sampled in the breeding range (data not shown). The estimates are independent of between-sex sampling biases, and so the observed deficit of females probably reflects within-sex behavioural differences, for example that individual females display a higher degree of preferences with respect to region and/or residence time in the breeding range than do males.

Our results demonstrate that genetic tagging is effective even in a large population of wide-ranging and inaccessible mammals such as cetaceans. Further, the data obtained from genetic tags can be used to address evolutionary[20], demographic[19] and behavioural[21]

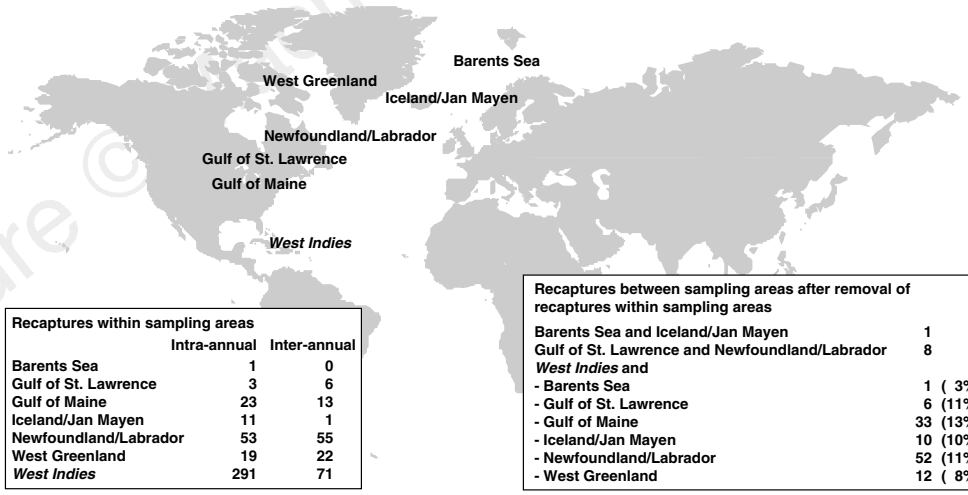
questions to which traditional tagging methods are unsuited. Because all eukaryotes possess microsatellites[22], individuals within any taxon can in principle be identified reliably from minute quantities of tissue. Such tissue is commonly derived from biopsies[23], but can also come from sloughed skin[5], shed hair[24] or fecal material[25,26], thus potentially allowing genotyping and individual recognition even of unobserved animals. However, the validity of a genetic tag depends on there being a sufficiently low probability of identity[27]. An underestimate of this probability can be caused by unrecognized population substructure or linkage disequilibrium among loci. Additional analyses addressing these issues, as well as checks with other data (such as the sex of recaptures, as in this study) must be performed to ensure the validity of the overall probability of identity. Similarly, a genetic tag consists of data from multiple loci, each of which are prone to handling errors in the laboratory. With proper laboratory procedures, such errors can be rendered minimal (here estimated to 0.0011 per locus), and so do not represent a serious obstacle to the detection of recaptures. Recaptures with an erroneous genotype will almost certainly be among the samples that match at all but one locus. If a sufficient number of loci has been analysed, none or few samples are expected to match at all but one locus, and so re-analysis of the discrepant locus presents only a minor additional effort. □

Methods

Expected number of samples with identical genotypes. The number of samples from different individuals with identical genotypes (across all loci) arising by chance was estimated from the probability of identity (I)[27]. No difference in the estimate was observed whether I was estimated from all samples of unique genotypes only. The expected number of samples from different individuals with identical genotypes arising by chance was first estimated separately within each sampling area, and subsequently between sampling areas (after removal of duplicate genotypes within each sampling area). The expected total number of such matches was estimated to be 0.32 and 0.27 within and between areas, respectively. No significant degree of linkage disequilibrium (which could cause an underestimate of I), nor any significant deviations from the expected Hardy–Weinberg proportions of genotypes, was observed after the removal of duplicate genotypes.

Expected number of matches between the Gulf of St Lawrence and Newfoundland/Labrador. The probability of observing six recaptures between years in the Gulf of St Lawrence was assessed by Monte Carlo simulations, under the assumption that the Gulf of St Lawrence and Newfoundland/Labrador constitute one intermixing feeding aggregation. Several tests, each of 1,000 simulations, were conducted over a range of the most likely abundance estimates derived from the data. Each simulation was conditioned on the number of recaptures observed in the sample and the order of sampling in the two areas. The probability of observing six or more individuals sampled more than once in the Gulf of St Lawrence (if part of the same feeding ground as Newfoundland/Labrador) was estimated to less than 0.0001.

Estimation of abundance and log-likelihood ratio test of equal numbers of males and females on the breeding range. Estimates of abundance and 95% confidence intervals were calculated as suggested previously[17]. In 1992, 382 males and 231 females were sampled on the breeding range; the corresponding numbers for 1993 were 408 and 265. Between the two years we observed 31 and 21 recaptures of males and females, respectively. The statistical significance of the difference in the estimated number of males and females was assessed with a likelihood-ratio test of the null hypothesis that the number of males was equal



Recaptures within sampling areas

	Intra-annual	Inter-annual
Barents Sea	1	0
Gulf of St. Lawrence	3	6
Gulf of Maine	23	13
Iceland/Jan Mayen	11	1
Newfoundland/Labrador	53	55
West Greenland	19	22
West Indies	291	71

Recaptures between sampling areas after removal of recaptures within sampling areas

Barents Sea and Iceland/Jan Mayen	1
Gulf of St. Lawrence and Newfoundland/Labrador	8
West Indies and	
- Barents Sea	1 (3%)
- Gulf of St. Lawrence	6 (11%)
- Gulf of Maine	33 (13%)
- Iceland/Jan Mayen	10 (10%)
- Newfoundland/Labrador	52 (11%)
- West Greenland	12 (8%)

Figure 1 Numbers and distribution of recaptures. Sampling areas in italics are in the breeding range. The percentage in parentheses denotes the proportion of recaptures relative to the total number of genotypes tagged in the given sampling area.

Table 1 Sample sizes, probabilities of identity, number of genotypes and individuals of each sex

Sampling area	Period	Samples	I ($\times10^{-7}$)*	95% confidence limits ($\times10^{-7}$)†	Genotypes	Males	Females
Barents Sea	1992–1993	36	8.46	4.9–38	35	13	22
Gulf of St Lawrence	1990–1995	65	1.94	1.26–5.52	56	28	28
Gulf of Maine	1990–1995	292	1.38	1.02–2.11	256	118	138
Iceland/Jan Mayen	1991–1993	112	1.43	0.88–3.28	100	50	50
Newfoundland/Labrador	1991–1995	572	1.34	1.07–1.86	464	237	227
West Greenland	1988–1994	189	1.23	0.89–2.10	148	75‡	72‡
West Indies (breeding range)	1989–1995	1,794	1.81	1.57–2.15	1,432	884‡	545‡
Without intra-area recaptures		2,491	1.51	1.34–1.75	2,368		
Unique genotypes only		2,368	1.51	1.32–1.72		1,331‡	1,033‡

* Probability of identical genotype[27] across all loci calculated from all samples (including recaptures).
† Estimated from 1,000 bootstrap samples.
‡ No sex was obtained for four samples.

to the number of females, assuming that the number of recaptures is hypergeometrically distributed. Asymptotically the test statistic ($-2\ln(\lambda)$, where λ is the likelihood ratio) is chi-square distributed with one degree of freedom, which implies that the probability of $-2\ln(\lambda)$ exceeding 4.14 is 0.042. Monte Carlo simulations (10,000 replicates) confirmed this probability.

Estimation of error rate. The error rate was estimated from the samples with identical genotypes at five but not six loci. From 2,368 genotypes just 64 such pairs (in all 117 unique genotypes) were detected, in 19 of which the individual from which the sample was collected had been identified by its natural markings[28]. Of these 19 incidences just 3 were from the same individual. Hence the 117 samples were estimated to include 9 samples with an incorrect genotype, which equals an error rate of 0.0011 per locus after inclusion of the 2 times 692 samples which match on all loci (presumably determined correctly). The upper limit of the 95% confidence interval was estimated to 0.0027 from Monte Carlo simulations, assuming a binomial distribution of errors, and a hypergeometric distribution of photographed whales.

Received 6 March; accepted 16 May 1997

1. Hammond, P. S., Mizroch, S. A. & Donovan, G. P. *Individual recognition of Cetaceans: Use of Photo-identification and Other Techniques to Estimate Population Parameters* (International Whaling Commission, Cambridge, 1990).
2. Tautz, D. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res.* **17,** 6463–6471 (1989).
3. Weber, J. L. & May, P. E. Abundant class of human DNA polymorphism which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.* **44,** 388–396 (1989).
4. Palsbøll, P. J., Larsen, F. & Sigurd Hansen, E. Sampling of skin biopsies from free-ranging large cetaceans in West Greenland: Development of new biopsy tips and bolt designs. *Rep. Int. Whaling Commiss. Spec. Iss.* **13,** 71–79 (1991).
5. Clapham, P. J., Palsbøll, P. J. & Mattila, D. K. High-energy behaviors in humpback whales as a source of sloughed skin for molecular analysis. *Mar. Mamm. Sci.* **9,** 213–220 (1993).
6. Maniatis, T., Fritsch, E. F. & Sambrock, J. *Molecular Cloning. A Laboratory Manual* (Cold Spring Harbour Laboratory Press, New York, 1982).
7. Bérubé, M. & Palsbøll, P. J. Identification of sex in Cetaceans by multiplexing with three *ZFX* and *ZFY* specific primers. *Mol. Ecol.* **5,** 283–287 (1996).
8. Clapham, P. J. & Palsbøll, P. J. Molecular analysis of paternity shows promiscous mating in female humpback whales (*Megaptera novaeangliae*, Borowski). *Proc. R. Soc. Lond. B* **264,** 95–98 (1997).
9. Palsbøll, P. J., Bérubé, M., Larsen, A. H. & Jørgensen, H. Primers for the amplification of tri- and tetramer microsatellite loci in cetaceans. *Mol. Ecol.* (in the press).
10. Katona, S. K. & Beard, J. A. Population size, migrations, and feeding aggregations of the humpback whale, *Megaptera novaeangliae*, in the western North Atlantic Ocean. *Rep. Int. Whaling Commiss. Spec. Iss.* **12,** 295–305 (1990).
11. Clapham, P. J. *et al.* Seasonal occurrence and annual return of humpback whales in the southern Gulf of Maine. *Can. J. Zool.* **71,** 440–443 (1993).
12. Palsbøll, P. J. *et al.* Distribution of mtDNA haplotypes in North Atlantic humpback whales: the influence of behaviour on population structure. *Mar. Ecol. Prog. Ser.* **116,** 1–10 (1995).
13. Larsen, A. H., Sigurjónsson, J., Øien, N., Vikingsson, G. & Palsbøll, P. J. Population genetic analysis of nuclear and mitochondrial loci in skin biopsies collected from Central and northeastern North Atlantic humpback whales (*Megaptera novaeangliae*): population identity and migratory destinations. *Proc. R. Soc. Lond. B* **263,** 1611–1618 (1996).
14. Stevick, P. T., Øien, N. & Mattila, D. K. Migration of a humpback whale (*Megaptera novaeangliae*) between Norway and the West Indies. *Mar. Mamm. Sci.* (in the press).
15. Clapham, P. J., Mattila, D. K. & Palsbøll, P. J. High-latitude-area composition of humpback whale groups in Samana Bay: further evidence for panmixis in the North Atlantic population. *Can. J. Zool.* **71,** 1065–1066 (1993).
16. Mattila, D. K., Clapham, P. J., Katona, S. K. & Stone, G. S. Population composition of humpback whales, *Megaptera novaeangliae*, on Silver Bank, 1984. *Can. J. Zool.* **67,** 281–285 (1989).
17. Seber, G. A. F. *The Estimation of Animal Abundance and Related Parameters* (Charles Griffin, London, 1982).
18. Barlow, J. & Clapham, P. J. A new birth-interval approach to estimating demographic parameters of humpback whales. *Ecology* **78,** 535–546 (1997).
19. Clapham, P. J., Bérubé, M. & Mattila, D. K. Sex ratio of the Gulf of Maine humpback whale population. *Mar. Mamm. Sci.* **11,** 227–231 (1995).
20. Roy, M. S., Geffen, E., Smith, D., Ostrander, E. A. & Wayne, R. K. Patterns of differentiation and hybridization in North American wolflike canids, revealed by analysis of microsatellite loci. *Mol. Biol. Evol.* **11,** 553–570 (1994).
21. Amos, B., Schlötterer, C. & Tautz, D. Social structure of pilot whales revealed by analytical DNA profiling. *Science* **260,** 670–672 (1993).
22. Tautz, D. & Renz, M. Simple sequences are ubiquitous repetitive components of eukaryote genomes. *Nucleic Acids Res.* **12,** 4127–4138 (1984).
23. Lambertsen, R. H. A biopsy system for large whales and its use for cytogenetics. *J. Mamm.* **68,** 443–445 (1987).
24. Morin, P. A. & Woodruff, D. S. in *Paternity in Primates: Genetic Tests and Theories* (eds Martin, R. D., Dixon, A. F. & Wickings, E. J.) 63–81 (Karger, Basel, 1992).
25. Constable, J. J., Packer, C., Collins, D. A. & Pusey, A. E. Nuclear DNA from primate dung. *Nature* **373,** 393 (1995).
26. Gerloff, U. *et al.* Amplification of hypervariable simple sequence repeats (microsatellites) from excremental DNA of wild living bonobos (*Pan paniscus*). *Mol. Ecol.* **4,** 515–518 (1995).
27. Paetkau, D. & Strobeck, C. Microsatellite analysis of genetic variation in black bear populations. *Mol. Ecol.* **3,** 489–495 (1994).
28. Katona, S. K. & Whitehead, H. P. Identifying humpback whales using their natural markings. *Polar Rec.* **20,** 439–444 (1981).

Acknowledgements. Most samples were collected during the international collaborative project (Year of the North Atlantic Humpback Wahle (YONAH). We thank T. H. Andersen, P. Arctander, C. Berchok, I. Bonnelly, M. Fredholm, J. Jensen, K. B. Pedensen, P. Raahaüge, J. Robbins, O. Vasquez and E. Widén for their support and assistance. Funds were obtained from the Commission for Scientific Research in Greenland, the Greenland Home Rule, the EU Biotechnology Program, the Danish and Norwegian Research Councils, the US National Marine Fisheries Service. The US National Fish and Wildlife Foundation, the Department of Fisheries and Oceans, the International Whaling Commission, the US State Department, the Aage V. Jensen Charity Foundation, the Dorr Foundation, the American–Scandinavian Foundation, the Exxon Corporation, and Feodor Pitcairn.

Correspondence and requests for materials should be addressed to P.J.P. (e-mail: pjpalsbo@uci.edu).

Neurodegeneration in Lurcher mice caused by mutation in δ2 glutamate receptor gene

Jian Zuo, Philip L. De Jager, Kanji A. Takahashi*, Weining Jiang, David J. Linden* & Nathaniel Heintz

Laboratory of Molecular Biology, Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA
** Department of Neuroscience, The Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205, USA*

Lurcher (*Lc*) is a spontaneous, semidominant mouse neurological mutation[1]. Heterozygous Lurcher mice (*Lc*/+) display ataxia as a result of a selective, cell-autonomous and apoptotic death of cerebellar Purkinje cells during postnatal development[2–4]. Homozygous Lurcher mice (*Lc*/*Lc*) die shortly after birth because of a massive loss of mid- and hindbrain neurons during late embryogenesis[5]. We have used positional cloning to identify the mutations responsible for neurodegeneration in two independent *Lc* alleles as G-to-A transitions that change a highly conserved alanine to a threonine residue in transmembrane domain III of the mouse δ2 glutamate receptor gene (GluRδ2). *Lc*/+ Purkinje cells have a very high membrane conductance and a depolarized resting potential, indicating the presence of a large, constitutive inward current. Expression of the mutant GluRδ2^{Lc} protein in *Xenopus* oocytes confirmed these results, demonstrating that *Lc* is inherited as a neurodegenerative disorder resulting from a gain-of-function mutation in a glutamate receptor gene. Thus the activation of apoptotic neuronal death in Lurcher mice may provide a physiologically relevant model for excitotoxic cell death.

The excitatory actions of glutamate in the central nervous system are mediated by several classes of glutamate receptors. Ionotropic glutamate receptors are ligand-gated cation channels that can be divided by their pharmacological properties into two classes: the NMDA (*N*-methyl-D-aspartate) and the AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazolepropionate)/kainate receptor types[6]. By sequence comparison, GluRδ2 is equally distantly related (~25% amino-acid identity) to both of these glutamate receptor subfamilies. It does not display glutamate binding or ion-channel activity when injected into *Xenopus* oocytes or expressed in mammalian cells either alone or in combination with other glutamate receptor subunits[7,8]. However, GluRδ2 null mutant mice lack cerebellar long term depression[9], a form of plasticity at the parallel fibre–Purkinje cell synapse that is thought to be important for some forms of motor learning.

We have designed an inter-subspecific backcross between *Lc* and CAST/Ei mice and established a detailed genetic and physical map over a three-megabase region of mouse chromosome 6 (Fig. 1)[10,11]. From the 504 informative meioses, four recombination events define the *Lc* locus, of which three are on the telomeric side and one on the centromeric side. The non-recombinant region identified by these recombination events is approximately 110 kilobases in length, and a physical contig consisting of four yeast artificial chromosomes (YACs) and five bacterial artificial chromosomes (BACs) covering this locus has been established[10,11]. A second *Lc* allele (Lc^J), which is phenotypically indistinguishable from *Lc*, has