Discovery of the Principal Cystic Fibrosis Mutation (F508del) in Ancient DNA from Iron Age Europeans

Philip Farrell[1], Cedric Le Marechal[2], Claude Ferec[2], Malika Siker[1], and Maria Teschler-Nicola[3]

[1] Departments of Population Health Sciences and Pediatrics

University of Wisconsin, Madison WI 53726, USA.

[2] Laboratoire de Génétique Moléculaire et Génétique Epidémiologique, INSERM U613,

Université de Bretagne Occidentale, 29220 Brest, France.

and the

[3] Department of Anthropology

University of Vienna and the Naturhistorisches Museum, A-1014 Vienna, Austria

To whom correspondence should be addressed. E-mail: pmfarrell@wisc.edu

The most common, life-threatening autosomal recessive disease of Europeans and Euro-Americans, cystic fibrosis (CF), occurs predominately in patients with the F508del mutation.[1] Although F508del is currently detectable as a single allele in 1/30-1/40 Europeans[2-4] and Euro-Americans,[5] it has not been determined what heterozygote selective advantage(s) might account for its relatively high prevalence. Indirect evidence[6] suggests that this mutation was present in Brittany at least 3000 years ago, but no direct analyses of ancient DNA have been reported to identify F508del and clarify its frequency in prehistoric inhabitants of Europe. Here we show that F508del was present in 3 of 32 Iron Age inhabitants of Austria from whom DNA could be recovered from molar teeth using procedures that fulfill authenticity criteria.[7] Because these individuals, who were buried in cemeteries along the Danube river, were shown by radiocarbon dating of isolated bone collagen to have lived there during 544-255 BC, this indicates that the F508del mutation is definitely more than 2000 years old and that CF (the disease) was present among them. More generally, the apparent enrichment of this Iron Age population in F508del suggests an evolutionary advantage in their environment that can be investigated by interdisciplinary strategies of paleoepidemiology.

Although more than 1500 mutations have been identified in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene, the F508del allele[1] accounts for about 70% of cystic fibrosis (CF) chromosomes in known patients and accounts for the relatively high incidence of the disease.[2, 3, 8] To address the gap in knowledge regarding potential selective heterozygote advantages proposed[7] to explain the prevalence of F508del-CF, we have initiated an interdisciplinary paleoepidemiology investigation focused on the hypothesis that exposures during the post-Neolithic population expansions/migrations[10] enriched tribal sub-populations in this mutant allele of the *CFTR* gene. Initially, we are seeking direct evidence demonstrating that CF was associated with prehistoric Europeans, namely the Celts who emerged as a distinct culture after the Bronze Age. The Celtic tribes as a warrior society with master metal workers then flourished during the Iron Age and, after large scale migrations,[10] occupied the regions of Europe where CF is most common.[2, 4, 8] The availability of relatively large populations of Iron Age skeletons provided an opportunity to determine if F508del was present in their DNA.

Iron Age cemeteries with Celtic burials in Eastern Austria with access to waterways such as the Danube river were the sources of our specimens. After careful excavation, the human remains were given to the Department of Anthropology at the Natural History Museum in Vienna. From three collections of Iron Age skeletons buried in cemeteries near Vienna [Mannersdorf/Leitha (30 km southeast), Pöttsching (50 km south), and Franzhausen (40km west)], we selected molar teeth (replaced with casts) from individuals with at least one intact tooth and 8 mm core femur specimens from all skeletons with at least one identifiable femur of adequate size. The teeth were individually packaged and transported to Brest, France for the molecular studies described below. The femur specimens were taken to other laboratories for analyses such as radiocarbon dating of isolated collagen by accelerator mass spectrometry

(AMS) at the Oxford Radiocarbon Accelerator Unit (http://c14.arch.ox.ac.uk/ams.html#prep).

Each skeleton had been examined and assigned a probable age and sex from its osseous features;

some were examined again after ancient DNA results became available. The estimated age of

these individuals ranged from approximately 3 to 80 years. The AMS-radiocarbon dating of

isolated collagen[11] from femurs revealed an average of 386 BC, with a range = 544-255 BC for

15 specimens analyzed.

  Processing and analysis of ancient DNA was organized to address recommendations for

ensuring authenticity.[7, 12] In planning this investigation, we decided that top priority should be

given to avoiding contamination with modern DNA. Thus, it was concluded that molar teeth

would be preferable to bones because of their protective encasement and the potential for

multiple specimens. The key strategy, however, was to perform replicate analyses with

independent technicians in two separate buildings employing rooms that had never been used for

any *CFTR* research. Multiple negative controls were also introduced at each step of the

procedure, as described by Bramanti et al.[13] Only two teeth were processed per week, in order to

avoid any cross contamination. After experiencing failure to recover DNA from partially

degraded or softened teeth, namely 7 molars from the Mannersdorf/Leitha, collection (Table 1),

we used only intact teeth from the Pöttsching (N=10) and Franzhausen cemeteries (N=29). DNA

analyses were performed by combining PCR amplication of exon 10 of the *CFTR* gene, which

contains F508del, with concomitant amplification of short tandem repeat (STR) loci.[14] This

permitted detection of any laboratory-related contamination with modern DNA, as well as sex

determination by assessing Amelogenin polymorphism.[15] The F508del allele was identified by

its characteristic 3 base pair (bp) deletion compared to the wild type locus (Figure 1) and

confirmatory sequencing.

As shown in Table 1, analyzing 46 teeth yielded recoverable DNA in 32 with none from 7 degraded teeth of Mannersdorf/Leitha collection, 9 of 10 from Pöttsching, and 23 of 29 from Franzhausen. Sex concordance was demonstrated in 27 of the 32 individuals, but 3 of 5 discordant skeletons had uncertain male/female assignments from osseous examinations. This finding adds to fulfillment of authenticity criteria.[7, 12] DNA fingerprinting of the 32 amplified extracts revealed 3 adults with F508del demonstrated by replicated analyses showing the characteristic 3bp deletion, as shown in Figure 1. There were two individuals with the F508del allele from Franzhausen and one from Pöttsching. The Franzhausen burials were at widely separated graves. As shown in Figure 1C, the presence of F508del was reconfirmed in duplicate with another tooth from Franzhausen #203 two-years after the original discovery. The 3bp deletion was confirmed by sequencing the PCR product. This provides unequivocal evidence that she (a 25-30 year-old female buried in 351 ± 30 BC) was a CF heterozygote.

Demonstrating F508del in these specimens indicates that CF was present in Iron Age European populations and that the mutation is at least 2300 years old. The age of the principal CF mutation has been matter of debate and uncertainty. Although Morral et al[16] suggested that F508del is 2600 generations (52,000 years) old, this indirect estimation has been criticized by Kaplan et al[17] who proposed 868 generations. In addition, Fichou et al[6] recently showed that in Western Europe (Brittany) the oldest ancestor of the F508del mutation was from ~3000 years ago.

The presence of the F508del-CF mutation in 3 of 32 ancient DNA extracts from these teeth might seem higher than expected, but enrichment of Iron Age European populations in this mutation is consistent with a heterozygote selective advantage. Although there has been much speculation about factors that might explain the probable evolutionary advantage of F508del,[9]

investigation of human populations to test the various hypotheses is sorely needed. We believe

that Iron Age skeletons abundant in Europe will permit such studies. Further paleoepidemiology

research on other ancient Celtic sub-populations[10] will also be needed to delineate F508del

frequency in ancient DNA throughout Europe[8] where the mutation proliferated over the centuries

in a geographically specific pattern[4, 8, 18] aligned with Celtic migrations.[8, 10] The methods that we

have established, particularly the use of molar teeth processed to avoid modern DNA

contamination, will facilitate such an investigation.

Methods

For ancient DNA analyses in two separated laboratories, we first cleaned selected molar teeth

with milliQ water and a scrapper, chemical decontamination was accomplished with one hour

incubation of a cleaning solution (1%SDS, 25mM NaEDTA, 20µg.ml^{-1} proteinase K) and

periodical wortexing. The solution was then removed, several rinsings with milliQ water were

performed, and the teeth were air-dried during one night. The teeth were then crushed in a

mineralogic mill to obtain a thin powder which was separated in three aliquots. Two of them

were subjected to immediate analysis and the last one was stored at -20°C. A decalcification

solution (EDTA 0.5M + 20µg.ml^{-1} proteinase K) was added to the powder and incubated

overnight at 37°C with agitation before decantation by centrifugation. One ml of the supernatant

was introduced in the DNA extraction column (Qiamp UltraVirus kit – Qiagen) which was

processed according to the manufacturer.

We quantified the concentration of the extracted DNA (quantifiler kit –

Appliedbiosytems, Foster City CA) in order to normalize the amount of DNA introduced in the

PCR (about 1ng). We decided to combine the PCR amplification of the CF gene exon 10

6

(containing F508del) with several STR sequences available in commercial kits in order to be sure

of the variability between teeth and to identify potential modern, laboratory-related DNA

contamination.[13] Thus, the STR sequences of the two technicians were determined and

monitored with each analysis. The first 30 teeth were analyzed with the SGM kit

(Appliedbiosytems, Foster City CA) and the last ones with Powerplex 16 kit (Promega, Madison

WI) which gave most robust results. Each aliquot was amplified at least twice, with one day

between each replicate. Sex was determined by assessing Amelogenin polymorphism with the

above kits. Short amplicons were preferentially detected and results interpreted after first

determining that the no template controls were negative and if at least three STRs were amplified

in parallel in the two aliquots. Whenever the F508del mutation was detected by its characteristic

3bp deletion compared to the 90bp of the wild type allele, this result was confirmed by

sequencing the PCR product (sequence data available upon request--see Supplementary

Information).

References

1.   Kerem, B., et al. Identification of the cystic fibrosis gene: Genetic analysis. *Science* 245, 1073-1080 (1989).

2.   Southern KW, et al. A survey of newborn screening for cystic fibrosis in Europe. *J Cyst Fibros*. 6, 57-65 (2007).

3.   Bobadilla, J.L. et al.   Cystic fibrosis: A worldwide analysis of *CFTR* mutations-correlation with incidence data and application to screening. *Human Mutation* 19:575-606 (2002).

4.   Lucotte G, Hazout S, De Braekeleer M. 1995. Complete Map of Cystic Fibrosis Mutation DF508 Frequencies in Western Europe and Correlation between Mutation Frequencies an Incidence of Disease. *Hum Biol* 67:797-803 (1995).

5.   Gregg RG, et al. The application of DNA analysis in a population screening program for neonatal diagnosis of cystic fibrosis: comparison of screening protocols. *Am J Human Genetics* 52, 616-626, (1993).

6.   Fichou Y, et al. Estimating the age of CFTR mutations predominantly found in Brittany (Western France). *J Cyst Fibros* In press—published online (2007).

7.   Hofreiter M, Serre D, Poinar HN, Kuch M, Pääbo. Ancient DNA. *Nature Reviews Genetics* 2, 353-359 (2001)

8.   De Braekeleer M., Daigneault J. Spatial distribution of the DF508 mutation in cystic fibrosis: A review. *J Human Biology* 64, 167-174 (1992).

9.   Romeo, G., Devoto, M., Galietta, L.V.J. Why is the cystic fibrosis gene so frequent? *Hum. Genet*. 84, 1-5 (1989).

10.   Cunliffe, B. The Ancient Celts. Penguin Books. (1997).

11.   Stafford, T.W., Jr., Hare, P.E., Currie, L.A., Jull, A.J.T. and Donahue, D. Accelerator radiocarbon dating at the molecular level. *J Archaeological Sciences* 18, 35-72 (1991).

12.   Cooper, A., Poinar, H.N. Ancient DNA: Do it right or not at all. *Science*; 289:1139 (2000).

13.   Bramanti B, Hummel S, Chiarelli B, Herrmann B. Ancient DNA analysis of the delta F508 mutation. *Human Biology* 75, 105-115 (2003).

14.   Butler JM. Genetics and genomics of core short tandem repeat loci used in human identity testing. *J Forensic Sci* 51, 253-265 (2006).

15.   Sullivan KM, Mannucci A, Kimpton CP, Gill P. A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene amelogenin. *BioTechniques* 15, 636-641 (1993).

16.   Morral, N., et al. The origin of the major cystic fibrosis mutation (delta F508) in European populations. *Nat Genet* 7, 169-175 (1994).

17.   Kaplan NL, Lewis PO, Weir BS. Age of the ΔF508 cystic fibrosis mutation. *Nature Genetics* 8, 216-217 (1994).

18.   European Working Group on CF Genetics. 1990. Gradient of distribution in Europe of the major CF mutation and its associated haplotypes. *Hum Genet* 85:436-45 (1990).

Acknowledgements

Author Information

The authors declare no competing financial interests. Correspondence and requests should be addressed Philip M. Farrell (pmfarrell@wisc.edu).

Table 1. Ancient DNA Results Using Analysis of Short Tandom Repeat Loci and CFTR Gene Exon 10

Cemetery	Samples available	Selected	DNA recovered	Gender concordance	del3bp (F508del)
Mannersdorf/Leitha	50	7	0		
Pöttsching	17	10	9	8	1
Franzhausen	52	29	23	19	2[+]

[+] tooth FRANZ203 was replicated and the 36p deletion confirmed by sequencing (See Figure 1)

Figure Legend

Figure 1.   This figure provides examples of three "fingerprints" with their corresponding

electrophoregrams of the blue channel (FAM labeled). The horizontal axis displays the size

(basepairs) of the PCR products. The vertical axis represents arbitrary fluorescence units. The

CFTR exon 10 amplicon was added to the nine microsatellites and the gender marker

Amelogenin of the SGM kit (Appliedbiosystems, Foster City CA) which allowed a simultaneous

analyzes. The labels below D3S1358 and vWA microsatellites indicate the number of repeats for

each allele. The 90bp peak corresponds to the wild type allele (solid arrow) of the *CFTR* exon

10, whereas the F508del allele was screened based on the deletion of 3bp (ie 87bp) (broken

arrow). Double peaks with 1bp difference (non template mediated 3' end A adduct) are a

systematic artifact due to the low DNA quality and additional PCR cycles (35 cycles). Sample

Franzhausen #74 (panel A from January 2007) is the wild type whereas replicates of the sample

#203 (panel B from February 2005 and C from January 2007) indicate an heterozygous F508del

genotype (alleles 87 and 90bp).

A

FH 74
(1/07)

B

FH 203
(2/05)

C

FH 203
(1/07)