

DIALECTOMETRY AND GENETIC DEMOGRAPHY OF  
CYSTIC FIBROSIS. WHEN RESULTS CONVERGE:  
THE CASE OF WESTERN BRITTANY<sup>1</sup>

*Nadine Pellen & Tanguy Sollicec*

**0. Introduction**

Luigi Cavalli-Sforza's research (Cavalli-Sforza et al., 1989; 1996), which underlined the convergent potential of genetics and linguistics for reassessing the human past, was initially welcomed with great enthusiasm within the framework of what was called the *New Synthesis* (Excoffier et al. 2002; Métoz 2006). Recent advances in genetics, such as the possibility of sequencing ancient DNA, have led to the point where we can now seriously envisage the study of our origins with increasing precision (Slon et al. 2018).

However, the results to date have been relatively disappointing since the convergences are less clear-cut than it was initially hoped (Barbieri et al. 2014; Ben Hamed & Darlu 2007; Nettle & Harriss 2003; Sims-Williams 1998). Precise matches between the linguistic and genetic data are neither clear-cut nor are they particularly frequent (Pakendorf 2014; Campbell 2015; McMahon 2004).

Yet, recent studies focusing on precise aspects of linguistic data in certain regions of the globe, such as onomastics, show that it is possible to identify interactions between linguistic facts and genetic findings (Manni & Barraï 2001; Manni, Heeringa & Nerbonne 2006; Longobardi et al. 2015).

This article was born out of two separate approaches carried out independently of one another. The first dealt with genetic demography and the second with variational linguistics, aiming to study the population and language of Brittany from a certain perspective.

In their recent study of the genetic history of France, Saint Pierre et al. show that Brittany "is substantially closer to the population from north-west Europe than to the north of France, in spite of both being equally geographically close" (2020: 863). They propose that the Bretons' earliest ancestors could be the descendants of early Neolithic pastoralist nomads from the Steppes (SP) who would have arrived in Brittany (i.e. the 'NW cluster') via north-western Europe.

---

1. We express our heartfelt thanks to Professor German for having contributed valuable linguistic and historical information to this article and for having translated it. It was he who first saw the connection between our respective research and introduced us to each other. It was at his suggestion that we wrote this joint article. Of course, any remaining errors are our own.

The highest level of SP population in the NW cluster may be either due to its position, at an extreme of the possible Neolithic expansion or to later, and more recent migration from northern Europe with high SP proportion, i.e., Celtic and/or Anglo-Saxon (ibid.).

The second hypothesis would support the idea of a more recent migration from northern Europe with high SP proportion, i.e., Celtic and/or Anglo-Saxon.<sup>2</sup>

Migration of Britons in what was at the time Armorica (and is now Brittany) may explain this closeness. These migrations may have been quite constant during centuries although a two-wave model is generally assumed. A first wave would have occurred in the 4<sup>th</sup> century when soldiers from British Isles were sent to Armorica, whereas the second wave consisted of Britons escaping the Anglo-Saxon invasions.<sup>3</sup> Additional analyses, on larger datasets may be required to discriminate between these various models (ibid.).

Our initial hypothesis is that the convergences in linguistics and genetics may be explained by a unique historical event, namely, the Brittonic settlement of the Armorican peninsula from the 4<sup>th</sup> through at least the 7<sup>th</sup> centuries and perhaps as late as the 8<sup>th</sup> century.

The data studied in the context of our respective research arises from the convergence of two very distinct sources. Sollicec's work was based on the dialectometric analysis of the data from the *Nouvel Atlas de la Basse-Bretagne (New Linguistic Atlas of Lower Brittany)*, Le Dû 2001, hereinafter NALBB), a linguistic atlas of the Breton language.

Pellen's work results from in-depth research on the genealogical database for cystic fibrosis patients in Brittany, the French region where the disease is most common (cf. map 2 below). The database was created by Professor Chaventré (INED, Paris) during the 1990s and was developed by Nadine Pellen between 2007 and 2019. Today, it contains 1,542 patients, more than 1,000 genealogical trees comprising more than 630,000 individuals having lived between the 16<sup>th</sup> century and the present day. The genetic material was associated with a genealogical approach in order to understand the social emergence of the disease.

2. The precise demographic importance of the Anglo-Saxons settlements in Britain is a matter for discussion between historians and archaeologists specialised in this period. Geneticists such as Oppenheimer (2006) have concluded that the contribution to the genetic pool of the population inhabiting Britain by incoming Old English speakers during the post-Roman period was relatively minor.

3. The idea of continuity in the Brittonic settlement patterns is also suggested by some historians, see Brett 2019. The Anglo-Saxons settlements seem to have had little influence (Coumert 2010).

In both cases, the results obtained by working with contemporary data make it possible to open a window onto the history of Brittany and to shed additional light on the origins of its people and language. Nevertheless, we wish to stress that we are fully conscious of the potential pitfalls in assuming links between language and culture as opposed to genetic backgrounds. This work on rare genetic variations, based on a very large set of genealogical data, and associated with a linguistic analysis using tens of thousands of phonetic dialectometric criteria, may confirm the study of Saint Pierre et al. which are derived from the analysis of many genes. St Pierre et al. (2020: 863) maintains that “exploration of more recent rare genetic variation will allow a better disentangling.” Such variation, to our mind, is the study of the *CFTR* gene.

In section 1 of this article, we present the data processed and analysed during our respective research. Then, in section 2, we present the methods we used and the results we obtained. Finally, in section 3, we propose a number of hypotheses in an attempt to provide a more solid historical interpretation for the ethnogenesis of the Breton population.

## 1. The data examined

### 1.1. Linguistic data

Solliec’s dialectometric approach was based on the NALBB which presents a large quantity of precise phonetic transcriptions for 600 lexical and morphological items recorded at 187 points within the Breton-speaking area of Western Brittany. Solliec studied the linguistic data extracted from the NALBB using linguistic distance measurements in the context of a dialectometric linguistic analysis which will be explained in greater detail below.

The Breton recorded in the NALBB reflects purely vernacular dialects or basilects. These micro-varieties have been transmitted orally from generation to generation and, as such, are viewed by the speakers as socially inherited. Their use of vernacular Breton has very rarely been expressed through the medium of writing even though certain varieties have been exposed to written forms of the language via what Le Dù and Le Berre (1996) refer to as the ‘ecclesiastical standard’, a register of Breton which could be described as an *interlangue*, that is, a local literary koinè intended especially for religious purposes. The modern standard Breton was recently created with the intention of providing a unified literary language for all Breton speakers and, for this reason, does not yet possess what might be described as ‘historical depth’.

Jean Le Dù’s initial objective was to continue and expand Pierre Le Roux’s *Atlas Linguistique de la Basse-Bretagne* (Le Roux 1924–1963, henceforth ALBB; Le Dù 1972). The NALBB, however, focused mainly on the phonetic and lexical dimensions

of Breton. The surveys started in 1969 in Saint-Rivoal in central Finistère and continued until 1997. The survey points chosen ranged from the island of Ouessant in the far west to the Rhuys peninsula in the southeast of Morbihan. The points of the NALBB where fieldwork was conducted is significantly more dense than that of the ALBB. The network selected for the surveys is particularly concentrated since, within the Breton-speaking area, close to one parish out of three is represented as can be seen on map 1.



Map 1. Survey points selected for the NALBB

This atlas also benefited from technological advances in that the surveys were all tape recorded. The audio tapes are now partially digitized and are accessible on the internet. Likewise, the maps were computer-generated but this was done before Unicode fonts were commonly in use.

One of the strengths of the NALBB is that the recordings were all transcribed by the author, who was a native speaker. Indeed, in order to limit the number of transcribers who would inevitably have varying perceptions of the pronunciations, Le Dû believed that having only a single phonological filter would grant greater



cohesion to the whole and thus avoid multiple phonetic interpretations of the data by different transcribers. In order to account for the pronunciations at each point as precisely as possible, he adopted a relatively fine phonetic notation. In all, around 110,000 forms were transcribed in the NALBB.

As part of Sollic's PhD dissertation (2021), he selected 106 maps (out of the NALBB's 600) limited to only one lexeme throughout the Breton-speaking zone. For instance, Map 310 'SEAU', meaning 'bucket' was excluded because two lexemes, *sailh* and *kelorn*, are used in various regions of the country. Analysing the same lexical item for all of Western Brittany guaranteed that the data would be statistically harmonious for each point of the atlas.

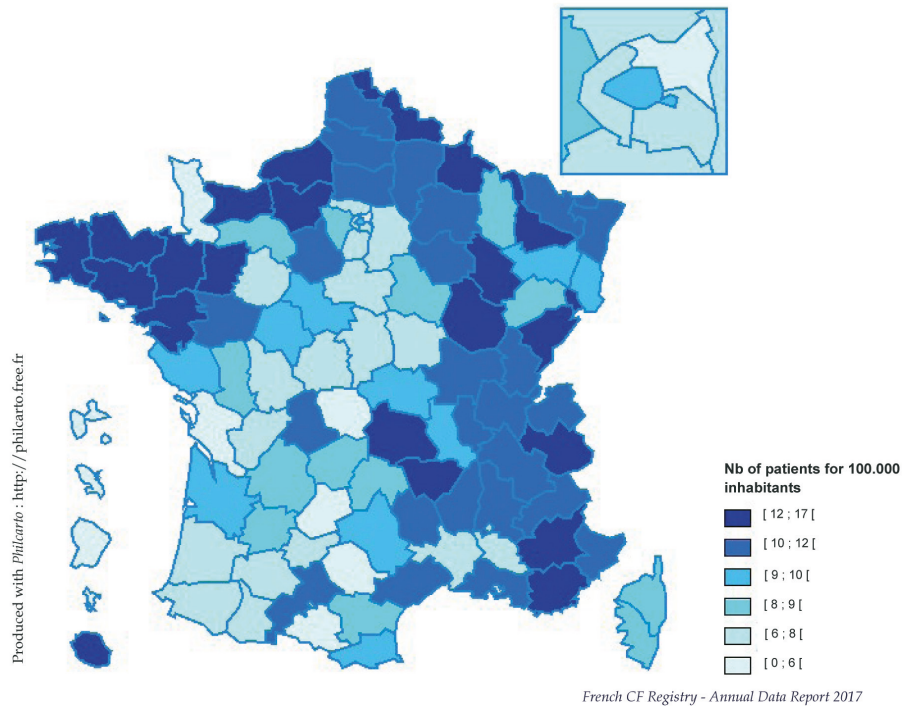
### 1.2. Genetic data

From a socio-genetic point of view, Pellen used data from a demographic study the objective of which was to explain the frequency and distribution of cystic fibrosis observed today in Brittany by cross-checking and comparing genetic and genealogical data.

Cystic fibrosis is a genetic disease that primarily affects patients' respiratory and digestive systems. It is said to have autosomal recessive inheritance. In other words, for a child to develop cystic fibrosis, each of their biological parents must themselves be carriers of a mutation that they have transmitted to their child. The parents are then called 'healthy carriers' or 'heterozygotes'. However, the child is 'homozygous', a carrier of two mutations and, therefore, an active carrier of the disease.

Cystic fibrosis is due to the alteration (mutation) of a gene called *CFTR* which is located on the long arm of chromosome 7 (i.e. 7q31). The *CFTR* gene responsible for the disease was identified in 1989 (Rommens et al. 1989; Riordan et al. 1989). More than 2,000 different mutations or modifications of the DNA sequence of this gene have been identified in the *CFTR* gene since its discovery in 1989. The F508del mutation is the most common in France, where it represents 83,4% of mutations (Bellis, Dehillotte & Lemonnier 2019). More than two million people are believed to be 'healthy carriers'. The *CFTR* gene contains the information to produce a protein also known as CFTR (Cystic Fibrosis Transmembrane Conductance Regulator). It is this protein which regulates the passage of chlorine across the cell membrane. Due to mutations in the *CFTR* gene, the protein is either absent or abnormal, meaning that chlorine cannot properly cross cell membranes.

Brittany is one of the regions that is most affected by cystic fibrosis in France (map 2).



Map 2. Prevalence of cystic fibrosis by department (number of patients per 100,000 inhabitants)  
(Bellis, Dehillotte & Lemonnier 2019: 10)

The information used in this article derives from the Cystic Fibrosis, Genetics and Genealogies (hereinafter CFGG) database which lists 1,542 patients suffering from cystic fibrosis, who were followed up or diagnosed in Brittany between 1964 and 2014 as well as their direct ancestors.<sup>4</sup>

We have traced most of the patients' ancestors back to at least the 17<sup>th</sup> century due to the survival of extremely rich and complete genealogical records throughout Brittany.

Pellen identified 784 patients in the database who had Breton ancestry and 1,037 as 'healthy carriers' of the *CFTR* gene. Only the genealogies of the 'healthy carriers' were taken into account. The study was thus not directly based on the patients themselves but on their parents, who were carriers of one *CFTR* gene mutation, and their

4. The Ildys Foundation (Brest-Roscoff, France) owns the database (Acte réglementaire, *Ouest-France*, 2 May 2004). The data collected was divided into civil (i.e. the official status of each individual), genetic and genealogical categories. Thus, the date and place of birth, marital status and death record were registered for each patient, as well as genetic mutations and microsatellites (i.e. haplotypes), if such information was available.

ancestors. The study population numbered a total of 541,374 direct ancestors. When all of the common ancestors going back to the 16<sup>th</sup> century were factored in, the number mounted to 1,554,848 individuals, the average genealogical depth was 15.7 generations, and the average number of ancestors per person was 1,177 individuals.

## 2. The methods used and results obtained

### 2.1. Dialectometry: A quantitative method for the analysis of dialect data

The NALBB data were exploited following a quantitative method known as dialectometry. This approach, which was first developed by Jean Séguy (1971, 1973a, 1973b), then by Hans Goebel in Salzburg (1982, 1983, 1984, 2010) aims to quantify the degree of resemblance (and thus dissimilarity) between the linguistic characteristics.

From a methodological perspective, the choice of a quantitative method made it possible to overcome several difficulties. For instance, the use of isoglosses to represent linguistic distinctions between the points would have led to an illegible, chaotic result from a visual perspective. In this sense, given that all the data is treated identically, dialectometry provides an objective method that prohibits the linguist from arbitrarily selecting a few aspects of linguistic data.

At the heart of Sollic's dialectometric approach is the Levenshtein algorithm (Levenshtein 1966), a tool that is commonly used in computer science and designed to compare two strings of characters and assess the distance between them in terms of similarity and dissimilarity. This technique was applied to dialectometry for the first time by John Nerbonne at the University of Groningen and was systematically applied by his pupil Heeringa (2004) whose work yielded positive results in the treatment of dialect data (Nerbonne & Heeringa 1997, 2010; Nerbonne, Heeringa & Kleiweg 1999). Within the framework of Sollic's work, he retained a version of the Levenshtein algorithm as modified by Guylaine Brun-Trigaud (2014).<sup>5</sup>

The Levenshtein algorithm, which is also known as 'edit distance', calculates the number of operations necessary to transform 'string of characters A' into 'string of characters B'. To do this, the algorithm uses three distinct operations: a) deleting a character, b) inserting a character and/or c) replacing one character with another. A value is assigned to each operation so that the distance between the two strings can be quantified as shown in the following example (table 1).

---

5. For an example of how the approach was applied cf. Brun-Trigaud, Sollic & Le Dû 2016; Sollic 2017, 2019.

**Table 1. Comparison of two strings according to the Levenshtein algorithm: Point 131( Kernével) ['jɛot] vs. point 133 (Nizon) ['jɛwot]<sup>6</sup>**

131 Kernével	'	j	ɛ		o	t
133 Nizon	'	j	ɛ <sub>ɣ</sub>	w	o	t
Nature of the operation			Replacement of ɛ by ɛ <sub>ɣ</sub>	Insertion of _w_	Replacement of o by o	
Number of operations			1	1	1	

The rate of similarity between the two strings is obtained by applying the following formula:

$$100 - (\text{number of differences} \times 100) / (\text{number of characters in the longest string})$$

In the present case, from the example presented in table 1 we obtain a similarity rate of 40%:

$$100 - (300/5) = 40 \text{ or } 40\% \text{ similarity}$$

The use of Levenshtein distance is symmetrical: the 40% similarity corresponds to 60% linguistic difference. The results obtained by exploiting the nature of the operations performed by the algorithm were presented in Sollicec (2017, 2019).

In order to draw systematic comparisons, Sollicec established a network of segments in which he compared each point of the atlas with each of those surrounding it. Thus, of the 187 points of the NALBB, he created a network of 534 segments according to the inter-punctual method (map 3).<sup>7</sup>

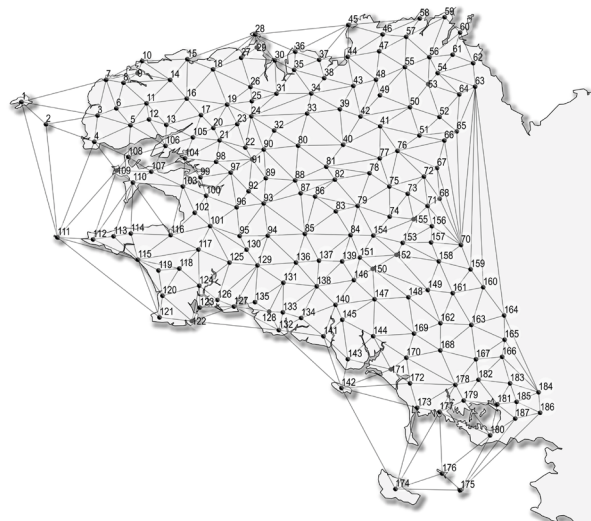
As in previous research on dialectometry (German 1984, 1993; Costaouec 2012 for Breton; Serva 2012; Serva et al. 2012 for Malagasy), Sollicec expected the geo-linguistic distribution of the similarity rates between points to highlight historical patterns that could then be identified and interpreted.

## 2.2. Results of the dialectometric analysis

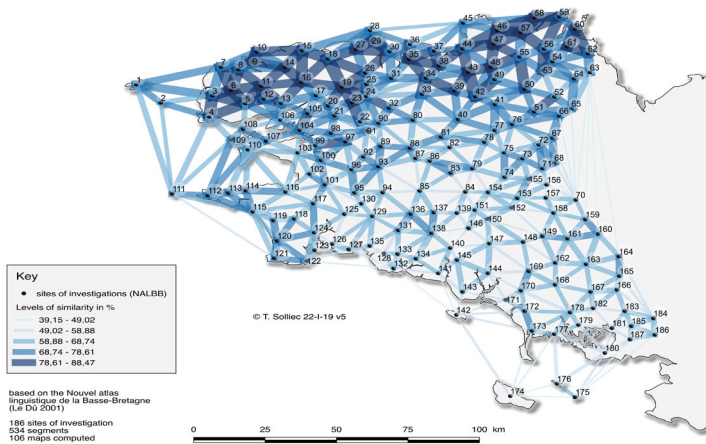
The results obtained by applying a dialectometric analysis to the linguistic data from the NALBB led to the following cartographic representation (map 4):

6. Both points are in southern Finistère (Cornouaille). The data are from map NALBB 176 'GRASS' geot, yeot.

7. The inter-punctual method allows the breakdown of the areal zone studied by linking each point of the atlas under study to its nearest neighbours. One thus determines an 'interpoint', that is to say, the relation between two points that permits the systematic comparison of the differences at each point of the atlas (Goebl 1983).



Map 3. Network of interpoints used to measure linguistic distance between points of the NALBB



Map 4. Distribution of linguistic similarity across the Breton-speaking zone

The average rate of linguistic similarity between the points of the NALBB was 68.43%. Results ranged from 39.15% for interpoint 63 through 160 (Pléguen – Naizin), to a maximum of 88.47% for interpoint 57–58 (Camlez – Plougrescant). The greater the similarity between two points, the thicker and darker the line connecting them. Table 2 provides an overview of the elements of descriptive statistics Sollic obtained following linguistic similarity measurements for the whole of Brittany.

The average rates referred to here are based on 106 NALBB maps (out of a total of 600) computed at 187 survey points of each map. The comparisons were calculated according to the phonetic composition of 106 words at each point of NALBB. In all, 19,716 forms were analysed (cf. section 1.1 above) giving a total of 56,604 comparisons.

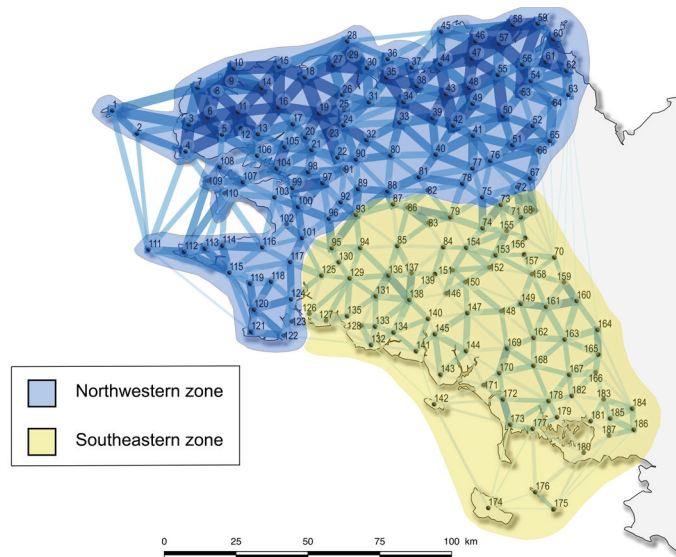
**Table 2. Results of descriptive statistics relating to the distribution of linguistic similarity in Breton-speaking zone**

Number of interpoints	534
Average	68.43%
Median	68.27%
Maximum	88.47%
3rd quartile	75.63%
1 <sup>st</sup> quartile	62.83%
Minimum	39.15%
Range	49.32
Standard deviation	9.52

A detailed analysis of the results and variation in the rates of linguistic similarity in areal space makes it possible to observe patterns of similarity and dissimilarity throughout western Brittany along a north-west / south-east axis in which two zones are clearly differentiated in several ways. A different colour has been assigned to each of them as can be seen in map 5.

The areal opposition between a north-western area and a south-eastern area is also counted as a ‘difference’ in the linguistic profile of each area. As can be seen in maps 5 and 6, the north-western area has highest rates of linguistic similarity and the variations from one inter-punctual segment to another are not very significant, indicating a relatively homogeneous linguistic space. The south-eastern

zone, however, groups together segments whose values are significantly lower and for which the variations from one segment to another are considerably greater.



Map 5. Division of Western Brittany in north-western and south-eastern linguistic zones<sup>8</sup>

Table 3 synthesises Sollicec's observations.

**Table 3. Statistical results with regard to the distribution of linguistic similarity in the north-western and south-eastern zones when compared to the whole of Breton-speaking Western Brittany**

	North-western zone	South-eastern zone	All of Western Brittany
Number of segments	323	211	534
Average rate of similarity	73.9%	60.04%	68.43%

8. After visual inspection of the results, the two zones were established on the basis of linguistic similarity rates. These rates were computed using the Levenshtein distance. The results presented in table 3 show that the profiles of these zones are markedly different.



DIALECTOMETRY AND THE GENETIC DEMOGRAPHY OF CYSTIC FIBROSIS

Median rate of similarity	74.1%	61.55%	68.27%
Maximum rate of similarity	88.47%	72.39%	88.47%
Minimum rate of similarity	51.58%	39.15%	39.15%
Ranges of measurements of similarity	36.89	33.24	49.32
Coefficients of variation	8.85	11.49	13.91

The profile of these two zones differ along several lines, as shown in figure 1:

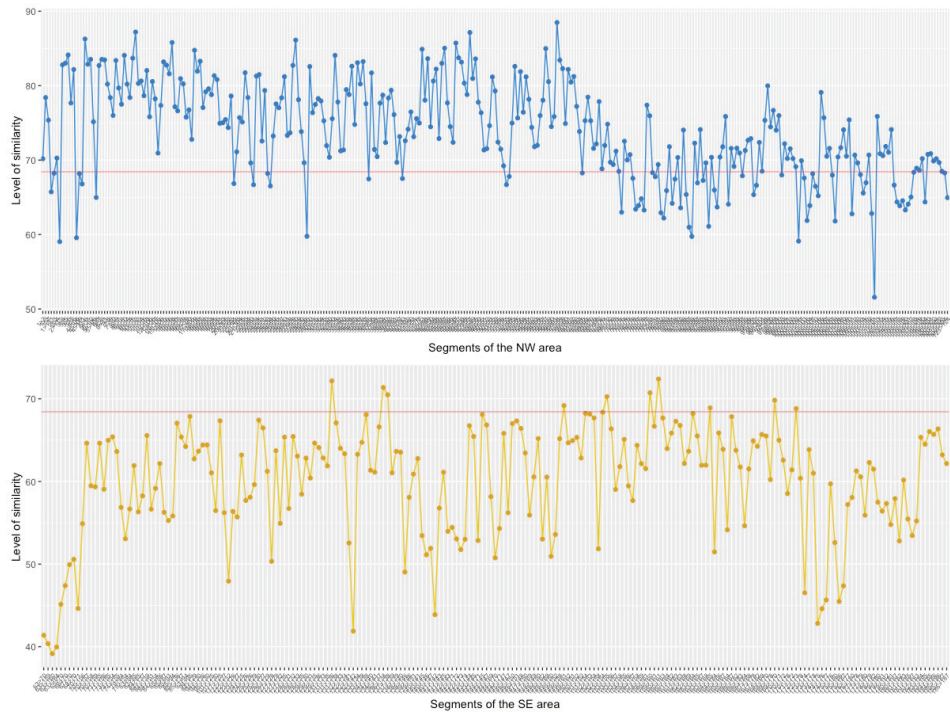
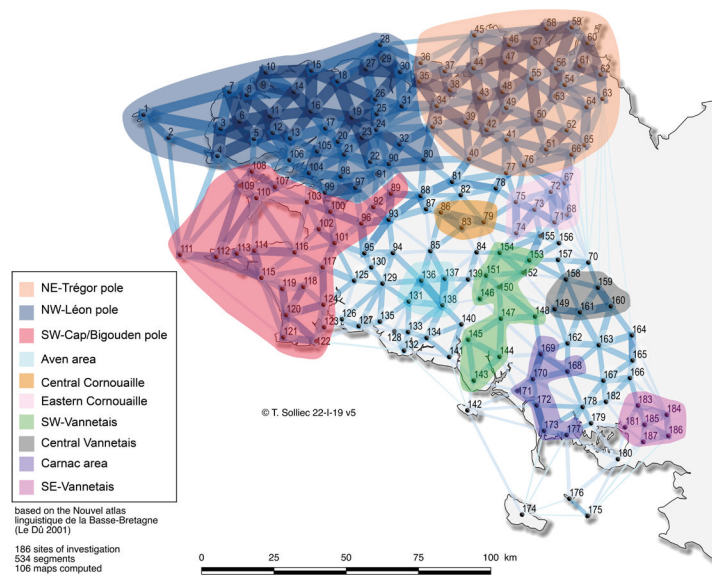


Figure 1. Linguistic similarity rate of interpoints according to the area



Several other facts can be observed regarding this north-western/south-eastern divide (in particular relating to the territories of the ancient Gaulish tribes: the *Ossismi* and the *Veneti*). Suffice it to say that the opposition between these two zones suggests that they were consolidated linguistically at different times (possibly after the year 300 AD) with different linguistic substrates. The general scholarly opinion is that the north-western zone of Brittany (corresponding to modern Léon and Trégor areas) corresponds to the region of Armorica settled by the Brythons (British Celts) from the 4<sup>th</sup> to the 7<sup>th</sup> centuries. The evidence from the south-eastern zone seems to mirror an older, more stable Gallo-Roman population which appears to have been affected by later western Brittonic adstratum influences. These differences in profile between the two areas suggests a chronological shift to which we shall return in section 3.

Furthermore, an examination of the internal linguistic structure of each of these two areas shows that several poles emerge within them, thus creating additional levels of analysis. Sollicie's dialectometric results identified 10 different major poles across Brittany which are represented in map 6.



Map 6. Distribution of the different poles of similarity across Brittany<sup>9</sup>

9. The subgroups were firstly identified by visual inspection of the maps, in an effort to identify the break lines in the north-western zone, and in the south-eastern zone, by isolating the groupings of segments that show a little more similarity than those in their immediate environment.

The north-western area thus consists of 3 comparatively similar poles in terms of linguistic distance thus reflecting comparative linguistic unity. On the other hand, we note seven distinct poles in the south-eastern zone revealing a far more fragmented linguistic landscape. Table 4 provides elements of descriptive statistics for each of these poles.<sup>10</sup>

**Table 4. Comparative table of statistical distinctions in terms of similarity**

Name of pole	Average rate of similarity	Maximum rate of similarity	Minimum rate of similarity	Standard deviation	Number of segments
NW-Léon	77.95%	87.19%	66.68%	4.65	101
NE-Trégor	77.44%	88.47%	66.69%	4.92	90
SW-Cap/Bigouden	68.22%	75.86%	51.57%	4.16	55
Aven area	69.92%	73.28%	66.6%	2.99	5
Central Cornouaille	69.99	72.27%	65.9%	2.99	3
Eastern Cornouaille	71.71%	77.85	68.48	2.68	11
SW-Vannetais	66.34%	68.37	62.83	1.65	17
Central Vannetais	66.55%	72.39	61.55	3.93	7
Carnac area	65.89%	69.83	62.57	2.59	8
SE-Vannetais	64.76%	66.36	62.16	1.56	7

The division between the north-western and south-eastern zones also affects these poles. The poles in the north-western zone therefore have a higher level of linguistic similarity (73.9%), which is visibly above the average for the whole of Brittany (68.43%). They form homogeneous groups which have a fairly large territorial footprint.

10. Despite the small number of interpoints for these poles, Sollicec has nevertheless given the values for comparison; yet, these values should be treated with caution.

The other seven poles in the south-eastern zone are clearly more spatially circumscribed and usually include only a few survey points (between 3 and 10). These poles are characterised by fairly low similarity rates compared to those in the north-western area. They stand out, however, because the average level of linguistic distance in this immediate area is even lower (60.04%).

The poles in the north-western area are further characterised by well-defined linguistic features, some of which may have ties to areas of the insular Brittonic-speaking regions (Cornwall and Wales) as shall be seen in section 4. The uneven distribution of these possible Brittonic features led Sollicc to consider that they emerged as a result of a founder effect in the wake of which a pool of linguistic features was constituted, giving rise to new Brittonic varieties. One can compare this with the appearance of Creole languages or colonial varieties of European languages from the 16<sup>th</sup> century onwards (Mufwene 2001).

The poles that have thus been identified do not seem to have all emerged in the same way. Indeed, a detailed examination of each pole could allow them to be divided into several categories.

**Table 5. Typology of the different poles of similarity across Western Brittany**

Category 1	NE-Trégor, NW-Léon, SW-Cap/Bigouden
Poles with a high average linguistic similarity rate, with a large geographic distribution and linguistically separate	
Category 2	Aven region, the Central Vannetais, SE-Vannetais
Poles whose average rate of linguistic similarity is fairly high, but whose geographic distribution is limited and linguistically clearly distinct	
Category 3	Central Cornouaille, the Eastern Cornouaille, SW-Vannetais and the area around Carnac
Poles whose average rate of linguistic similarity is low, whose geographic extent is limited and where linguistic correspondences between them are weak	

**2.3. Understanding the socio-demographic context of a genetic disease**

As it has been already mentioned in section 1.2, Pellen initially reconstituted the genealogies while seeking to identify the ascendants of each patient in direct lines.

The collaterals<sup>11</sup> were not taken into consideration. This approach allowed her to reconstruct the ascending genealogies for the 1,037 Breton patients identified above.

Pellen converted the scientific data about the ancestors of the patients that were present in the CFGG database to a *Granite-muco* software created to analyse these data. In addition, in order to specify the family links between the members of the database, it allowed her, among other things, to make groupings according to mutation types and then make maps of their geographic distribution. Descriptive statistics were then automated to explain the corpus and dealt primarily with the completeness of the genealogical records, their genealogical depth, the age of individuals at the time of marital unions as well as preferential unions, etc. Algorithms integrated into this software also made it possible to locate, in time and in space, the individuals whose descendants include the largest number of patients or healthy carriers. *Granite-muco* also took into account the measurements of relatedness and even inbreeding between individuals.

Finally, the results obtained were mapped so as to define the geographical locations of the ancestors and thus pinpoint and verify the possible existence of *foci* for the different mutations. The *foci*, also called 'clusters', are spatio-temporal aggregates which effectively group together, across time and space, the ancestors of current patients.

The next step was to look for potential historical events or phenomena which may have shaped the data Pellen gathered to explain the genetic reality of the current patient population by overlapping their genetic and genealogical data. In this way, Pellen was able to assess certain demographic and social practices such as differential areas of establishment (i.e. a particularly strong concentration of the F508del mutation NW-Léon and NE-Trégor and a great density of the 1078delT mutation in Aven, Cornouaille), preferential marital alliances, privileged migratory flows and, very significantly, the near absence of close consanguinity for the generations studied which Pellen attributed to strict control of the population by the Catholic Church.

In fact, the family genealogies that were thus reconstructed fed the chronicle of social and genetic reproduction, particularly through the types of the marital alliances. In addition, it was possible to look at the genetic and social background of single individuals in a collective light. One may recall the following observation in respect of the origin of hereditary diseases in Quebec:

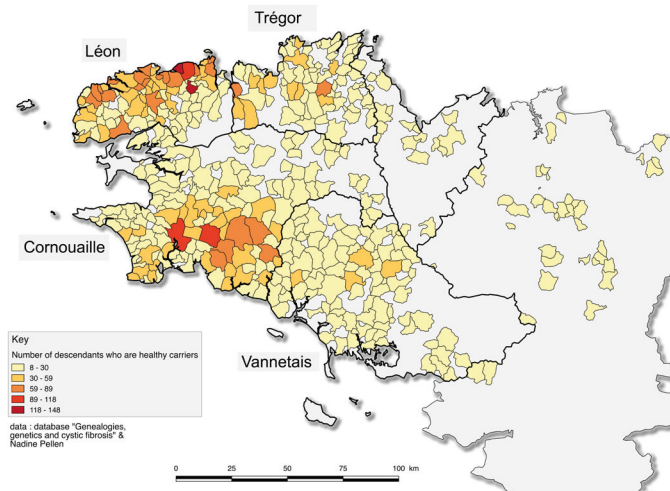
This problem has a collective character insofar as it is inscribed in the genetic heritage of a population. The genes that cause it are very common; they are passed down from one generation to the next and they will still be present in five, ten or twenty generations from now (Bouchard & Braekeleer 1992: 34).

11. Collaterals are the parents of a person who is not part of the patient's direct line of ancestry: brothers and sisters, uncles and aunts and their descendants, and cousins.

For this reason, the study of a genetic marker, like that of mutations in the *CFTR* gene in a defined population, brings us back to both the study of biological facts and to social and historical phenomena linked to movements of populations and their social structures.

## 2.4. Results of genetic data analysis

The study of genealogical and genetic data makes it possible to identify specific population basins whose geographical location is well defined. Generally speaking, the distribution of the ancestors of cystic fibrosis patients reveals significantly different locations across Brittany with the north-western part of Brittany showing a very concentrated number of people affected by the illness, and the extreme south-eastern zone in the Vannetais country, which is distinctly less affected by the disease (map 7).



Map 7. Location of healthy carrier descendants from couples appearing in the genealogy of patients present in the CFGG database

We can identify a heavily affected population living in Léon. Another focus of the disease is located in Cornouaille and a third population pool is located in the Trégor area. These regions correspond quite closely to the limits of the former bishoprics — Léon, Cornouaille, Trégor, Vannes — which have their origins in the early Middle Ages, the frontiers of which are believed to correspond to former Roman provincial boundaries. The latter, in turn, are thought to correspond to Gaulish tribal divisions (cf. German 1993). As we shall see, these three zones correspond very closely to those areas identified by Sollic as NW-Léon, NE-Trégor, SW-Cap/

Bigouden and the Aven area east of Quimper (table 4).

*CFTR* gene coding for cystic fibrosis and genealogical research highlighted the territories where the ancestors of the current patients lived. The results obtained reveal the cartographic representations seen in the annex 1. The data collected through civil status records since the 16<sup>th</sup> century testify to a very static population with strong geographic stability. The maps illustrating the foyers of the cystic fibrosis gene mutations (cf. table 7 below for the approximate dates of their implantation in Brittany) show that these communities have been rooted in these regions for at least a thousand years and, for some mutations, since at least 1,000 BC–1,750 BC. Only since the 19<sup>th</sup> century and, especially, since World War I, has the Breton population migrated to other regions of France on a massive scale. Each map of the annex 1 presents a characteristic distribution of the patients' ancestors locations. There are three main *foci*:

- a) the population settled in Leon is essentially that of the ancestors of the current carriers of the F508del, G551D and 1717–1G> A mutations. The F508del mutation is that which one finds most in the world. G551D, also known as the 'Celtic mutation', is more common in north-western Europe and, in particular, in Ireland (8.1%) and in United Kingdom (3%) and in Brittany (3.6%) (ECFSPR, 2017; Bobadilla JL, 2002). 1717–1G> A is a rare and relatively dispersed mutation (Calafell et al. 2004).
- b) In southern Cornouaille, to the west and particularly to the east of Quimper (Pays Bigouden/Cap and Aven area respectively), there is an important concentration of carriers of the 1078delT, W846X and 4005 + 1G> A mutations. In France, the 1078delT mutation in particular has only been found in patients of Breton origin (Scotet et al. 2003), more specifically from the Elliant area of the Aven pole. Very significantly for our argument, in Europe, it is found mainly in Wales but also in the Northwest Midlands of England near the Welsh border (Calafell et al. 2004). Note also that the F508del found in Léon is also present in the Aven area. According to the same source, the W846X and 4005 + 1G> A mutations were only found in the west of France.
- c) A third basin is located in Trégor, where we find the ancestors of contemporary carriers of the I507del mutation. The distribution of the I507del mutation is widely spread across Europe. In Brittany, there are two pools of very circumscribed ancestors, one in Trégor and another in eastern Brittany.

There is also a concentration of ancestors of carriers of the G551D and N1303 mutation in southern Morbihan but what is striking about this area is the relative absence of the disease.

Situated geographically at the tip of a relatively inaccessible peninsula, this

particular distribution and frequency of the *CFTR* gene mutations in Brittany has been increased by high inertia of the Breton population. Endogamy, has favoured this social inertia and had direct consequences on the distribution, structure and heterogeneity of the genetic pool.

The choice of spouses is also of considerable importance yielding important demographic, genetic, sociological, economic and even linguistic information. A reduced marriage market is another factor. Indeed, the choice of a spouse within one's parish was the norm in our corpus from the 15<sup>th</sup> to the 20<sup>th</sup> centuries, the densest exchanges being made within an extremely reduced radius: on average from 5 to 10 kilometres. As a result, the possibilities of unions were limited and depended on the place of residence, the number of singles from which relatives were genetically distant, the age of the applicants as well as their economic and professional status... Under such conditions, the choice of a spouse remained very restricted. But its consequences were invariable: they maintained the existing social structure and, as a corollary, the genetic structure of the population (Pellen 2015).

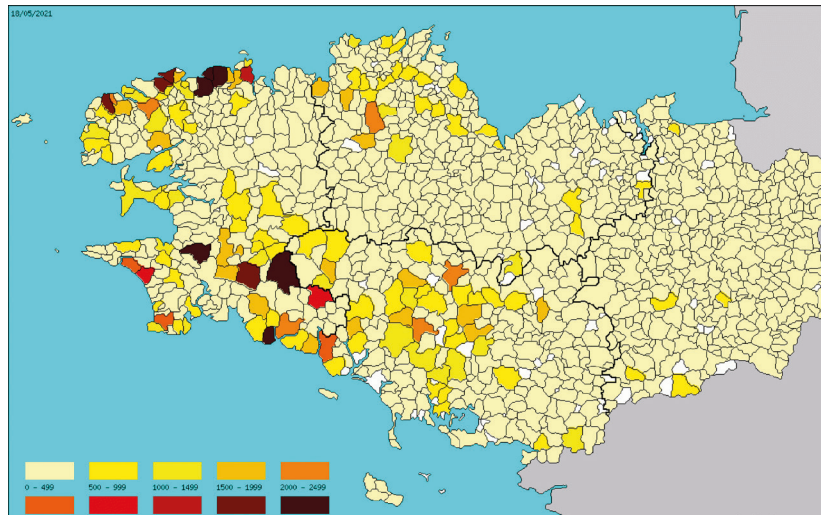
As a result, the populations studied are likely to be the descendants of the groups which were living in the same regions during the early Middle Ages. Indeed, the massive presence of the ancient F508del mutation suggests that the population has been implanted here for a long time (cf. table 7 below for the estimated dates of the mutations).

Maps 8 and 9 below show the number of people born, married and deceased in the same parish, as well as the parishes where the marriages took place, i.e. marriages of couples having the greatest number of carriers of the *CFTR* gene mutation responsible for cystic fibrosis among their descendants.

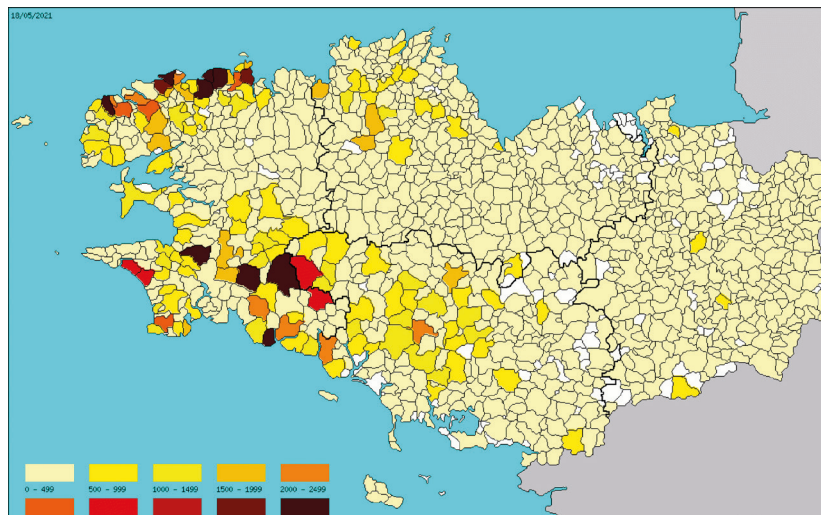
The similarities between these two maps represent, on the one hand, the geographically static profiles of the inhabitants (cf. map 8) and, on the other, specific geographic concentrations of the ancestors of the carriers of the *CFTR* gene mutations (cf. map 9) illustrating the corollary between regional roots and the absence of genetic mixing.

In order to test this, Pellen analysed inbreeding in patients over four generations. The conclusion is that there is no close consanguinity. Among the ancestors, very little evidence of this was found on account of the strict prohibition of such unions by the Catholic Church.





Map 8. People born, married and deceased in the same parish



Map 9. Location of weddings where couples had been related to several carriers



Only 858 ecclesiastical exemptions were granted for a total of more than 290,000 marital unions recorded.<sup>12</sup>

Located in very specific regions, the reconstructed family genealogies feed the chronicle of genetic reproduction, in particular through the forms of marital alliances in an endogamous environment (Segalen 1985). This firm local rooting also has had the obvious consequence of reducing the genetic diversity, absence of genetic mixing and an increasing prevalence of mutations of the *CFTR* gene.

### 2.5. Preliminary conclusions regarding the data analysis

The correspondence between genetic data presented by Pellen and dialectometric data presented by Sollicie demonstrate that the geographical distribution of the genetic pools of gene mutations corresponds closely to the pools of linguistic features located in the regions identified above. For this reason, both authors have concluded that the concordance of the genetic and linguistic data may result from a founder effect, the nature of which will be discussed in the next section.

## 3. Interpretation of the data

The methods that we used have led to surprisingly convergent results with the ones achieved by the *New Synthesis* school. We shall first detail these points of convergence before proceeding to a discussion of their possible causes.

### 3.1. Convergences between the genetic and linguistic data

The examination of the maps showing the distribution of common ancestors highlights a division between Western (Breton-speaking) Brittany and Eastern (Gallo-speaking) Brittany in terms of the prevalence of cystic fibrosis and, the number of ‘healthy carriers’, which is higher in the west of the Brittany than in the east (Pellen 2015: 109).

As we have already stated, this division can be equally observed in the distribution of the linguistic distance (with an exception of a few points in north-eastern Brittany along the Rance and the Couesnon rivers).

In addition, quite subtle correspondences were observed between several poles of distribution of the linguistic distance and certain genetic pools belonging to categories 1 (NW-Léon, NE-Trégor, SW-Cap/Bigouden) and 2 (Aven area, Central Vannetais, SE-Vannetais). These can be observed in annex 2.

---

12. As a result of the Council of Trent, from 1563 and until 1917, a union between two individuals with common ancestors requires a dispensation up to the fourth degree all inclusive. This corresponds to marriages with a common *trisaïeul* or great-great grandparent, in other words, the great-grandparents being brother and sister or the grandparents being first cousins. On the other hand, there are many distant relatives (going back at least 7 generations). It is these distant relatives that could have been decisive in the expansion of the *CFTR* genes in Brittany, even more so in Finistère where CF already featured in the population.

Table 6 provides a detailed summary of the geographical correspondences we arrived at after having annex 2 analysed.

**Table 6: Correspondences between the *CFTR* gene mutations and poles of distribution of the linguistic distance (+ correspondence; - no correspondence; ? insufficient evidence)<sup>13</sup>**

Type of <i>CFTR</i> gene mutation	NW-Léon	NE-Trégor	SW-Cap/Bigouden	Aven area	Central Vannetais	SE-Vannetais
F508del	+	+	+	+	?	-
G551D	+	-	-	+	?	-
1078delT	-	+	+	+	-	?
1717-1G>A	+	+	+	+	-	-
W846XelT	+	+	?	+	+	-
N1303KelT	?	+	+	-	+	+
4005+1G>A	+	+	+	+	-	-
G542X	+	+	+	-	-	-
I507d	-	+	-	-	-	-
G149R	+	?	+	+	-	-
621+1G>T	-	-	-	+	-	-
Number of correspondences	7	8	7	8	2	1

13. The connections are particularly marked for the poles of the north-west zone as well as in the Aven pole (see map 6, pole in light blue) which has a particular profile among the poles of the south-eastern zone and may be explained by a late Brittonic settlement in the region from Wales (cf. annex 2). We shall return to this below.

Close correspondences are displayed for the three poles in the north-western part of Western Brittany and include the Aven area. These poles all feature a high level of linguistic similarity and, at the geographical level, they also correlate very well with the distribution of 7 or 8 mutations out of the 11 *CFTR* gene under investigation. The level of correlation is far weaker in the south-eastern part of the country, namely, the Vannetais region. Here, only 1 or 2 correspondences can be observed between the linguistic data and the distribution of the genetic variants, chiefly in Central Vannetais and in the SE-Vannetais.

This concordance between the poles where linguistic similarity is the highest and where the ‘healthy carriers’ for each mutation are concentrated indicates the presence of genetic pools that have been well anchored geographically for many generations. In this regard, there is a marked difference between the north-western region (that includes the Aven pole, i.e. south-central and south-eastern Cornouaille) in which these pools are clearly identified, and the south-eastern zone corresponding to the Vannetais country. Here the correspondence between the genetic pools and the poles of the linguistic distance distribution is registered on a far lower scale.

These distinct regional profiles may very well stem from centuries-long Brittonic settlement which we shall discuss in section 3.2. Traditionally, scholars have explained such profiles on the basis of the linguistic differences between the NW-Léon, NE-Trégor and central Cornouaille dialects of Breton as opposed to the SE Vannetais dialect. We propose to take the regions of the former bishoprics into account, and this perspective brings us to the following set of observations.

### **3.2. The convergence between linguistic data and genetic heritage: the results of a two-faceted founding effect?**

The portrait drawn by the convergences between genetic data and distribution of the linguistic distance in Breton suggest a possible population movement the contemporary historical traces of which are very sketchy, namely the Breton migrations of the early Middle Ages (cf. table 7, dates of mutations). The convergences that we have noted between genetic material and linguistic facts underline a phenomenon which appears to stem from this migratory episode known as the ‘founder effect’. The results obtained allow us to consider it from the angles of population genetics and linguistics.

The notion of a founder effect stems from work in evolutionary biology and population genetics (Mayr 1942, 1954; Provine 2004). It describes the process by which the genetic diversity of a given population is less than that of a reference group from which this population comes. The founder effect is often associated with separation phenomena between such groups.

Whether it is viewed from a genetics or linguistic perspective, the ‘founder effect’ describes a similar process involving the reorganization of the distribution of the different linguistic or genetic traits in the new populations. On the one hand, this concerns all

the genes of a human group and, on the other, the different linguistic traits characteristic of the languages or varieties spoken by a given population of the same territory. The restructuring of genetic and linguistic phenomena is metaphorically described occurring in a 'pool of traits'. When introduced to a new territory, some of the latter take on new prominence and, over time, ultimately become characteristic of this new population.

The loss of contact between the source and emigrant groups is one of the causes of the 'founder effect'. The population which undergoes such a separation "would be the ideal place for a drastic genetic reorganization of the gene pool in the absence of any noticeable gene flow and under conditions of a more or less strikingly different physical and biotic environment" (Mayr 1982: 602).

The genetic data studied through the prism of cystic fibrosis in Brittany refers to the 'founder effect' resulting from a historical episode (cf. table 7 for the estimated dates of the CF mutations in Brittany). As soon as the biochemical mutation was introduced to this new territory, it was the social behaviour which was reproduced generation after generation that determined the diffusion (or geographical concentration) of this gene in the host population.

In our study, we are referring to source population in Britain during the early Middle Ages, and the Brython colonists originating from the insular communities who would have settled in the Armorican peninsula between the 5<sup>th</sup> and 7<sup>th</sup> centuries AD. At the time of their arrival, Armorica was primarily inhabited by the Gauls.<sup>14</sup>

It is the clustering of individuals sharing the same allele in the same geographical space that gives birth to a genetic pool. The genetic pool of a population, reflecting the diversity of its origins and history, is thus strongly determined by its demographic dynamics. A small migratory radius, strong homogamy and a population with many genealogical links perfectly illustrate the structuring process of the population into a stable nucleus. The ancestors of the current Breton population have, in fact, experienced similar physical and cultural constraints which explain this phenomenon. This stability gave rise to the linguistic poles identified in Map 6.

The local settlement had the consequence of reducing genetic diversity by limiting possible external contributions. The endogamy, more so than consanguineous marriages, tends to perpetuate the degree of genetic homogeneity peculiar to this population as a whole and to maintain the risk of certain genetic diseases at a relatively high level until the present day. This explains why the 'founder effect' of cystic fibrosis can still be felt nowadays.

Likewise, the outline of the different poles of the linguistic distance and their

---

14. The question of the languages spoken in Armorica at that time is a complex issue and still under discussion. Gaulish most probably survived as the common language of the peasantry, a form of Vulgar or Classical Latin was probably dominant among merchants, the ruling classes and clerics, depending on the area. Nevertheless, the dearth of sources for this period prevents us from having a clear picture of the linguistic situation during this period.

respective characteristics is the outcome of an analogous process. The concept of the ‘founder effect’ can be invoked to explain the emergence and the maintenance of these linguistic features.

### 3.3. A new source to better understand the history of Brittany?

In view of the sparse historical documentation for the Brittonic settlement period, our independent and separate analyses of contemporary linguistic and genetic data will hopefully shed valuable new light on the historical past of the Breton peninsula.

Indeed, we both consider that both the genetic and linguistic ‘founder effects’ most likely go back to the migration of populations from the British Isles to the continent and particularly to the Armorican peninsula between the 4<sup>th</sup> and the 7<sup>th</sup> centuries. It is during this period that the region came to be designated by the term *Britannia* or *Britanniae* in the 6<sup>th</sup> century by Grégoire de Tours (Coumert 2010: 15) rather than *Armorica*.

Despite this terminological evolution, as Coumert reminds us, “the birth of continental Brittany is a mystery” (ibid.). Indeed, in spite of the important work of compilation of the old sources of different kinds carried out by Fleuriot (1980),<sup>15</sup> there are no contemporary texts providing precise details about these settlements which would make it possible to understand the reasons and the exact course of events which led to this population movement. With the exception of a few metaphorical lines by the Brittonic cleric, Gildas (c. 490 – c. 570 AD), that he penned in his work *De excidio et conquestu Britanniae* (c. 540 AD), little is known:

Others made for lands beyond the sea; beneath the swelling sails they loudly wailed, singing a psalm that took the place of a shanty: ‘You have given us like sheep for eating and scattered us among the heathen.’

(Winterbottom 1978: 27–28)

This historical episode is known above all in later sources (e.g. the lives of saints of the 9<sup>th</sup> century in particular) and by the consequences of this migratory episode which are linguistic practices, anthroponyms (German 2006, 2014), toponyms or certain social uses. Following Fleuriot (1980), it was considered that these migrations took place in two stages, a so-called military migration in the 4<sup>th</sup> and 5<sup>th</sup> centuries followed by a movement of civilians supervised by civil and religious authorities, united by family links during the 6<sup>th</sup> century (Chédeville & Guillotel 1984).

However, this two-phase interpretation has recently been vigorously

---

15. Recently Coumert and Bouget have questioned the status of the sources collected by Fleuriot to write his book *Les origines de la Bretagne* (1980), pointing out that the additional volume he announced was never written (Bouget & Coumert 2019).

challenged, stressing that the material traces of these migrations remain imperceptible from an archaeological point of view (Brett 2011, 2019). Recent studies place the Brittonic immigration to the continent in the general context of the political and territorial reorganisation of Western Europe during the final collapse of the Roman Empire. However, many unanswered questions remain. Foremost among them are the social and geopolitical factors that prompted the Brythonic-speaking populations to migrate to Armorica and the linguistic effects of their settlement. Did the settlers impose their Brythonic varieties on the indigenous Armorican population as is generally supposed or, conversely, could the settlers' language been influenced by the Gaulish and Latin-speaking Armoricans?<sup>16</sup>

It appears as though the Brittonic migrations were part of a far more complex process than a straightforward two-phased migration. Several developments could thus have taken place simultaneously. Brett (2019: 21) reports that small groups of individuals may have been able to participate in this movement independently or under the protection of elites.

Taking this fresh approach to the analysis of the historical period into consideration, the examination of the genetic material makes it possible to underline certain links with the insular Brittonic-speaking world. For example, it is known that the 1078delT mutation in the *CFTR* gene that causes cystic fibrosis is found only in Wales (Cheadle et al. 1993), and in patients in south-western Brittany (Estivill, Bancells & Ramos 1997) with its hub around Elliant (in the centre of Aven Pole). This link testifies to a clear bond between the two populations which dates to the crucial period of Brittonic migrations, as we can see in table 7 below with the estimated age of the gene mutations.

The proximity of genetic data (mutations G551D, R117H, G542X, 1078delT, see table 7) reinforces the hypothesis of a common origin between the Brythons established in Armorica and the populations across the Channel. In the case of Brittany, the peculiarities of this genetic heritage appear to be the fruit of a founding effect and genetic drift (Scotet et al. 2003). What these researchers argue is that, when the population size is limited, a process of drift is instituted which leads, over time, to a progressive reduction in genetic diversity at the allelic level. The regional disparity observed in the ancestors of the present illness could be explained by the different geographical origins of the migrant groups: in particular, Wales, Devon and Cornwall. In the Breton context, as indicated previously, the current distribution and frequency of mutations of the *CFTR* gene in Brittany have been accentuated by the high geographical inertia of the local populations.

On the other hand, the appearance of various mutations of the *CFTR* gene in Brittany has been dated by geneticists (Fichou et al. 2008: 168–73) as indicated in table 7.

16. Cf. Guy 2014 for a sociolinguistic model and examples explaining how these linguistic shifts could have occurred.

**Table 7. Date of introduction of the various mutations of the *CFTR* gene in Brittany, their respective frequency and correspondences within Brittany and Ireland, Scotland and Wales<sup>17</sup>**

<i>CFTR</i> Gene mutations	Estimated date of entry in Brittany	Percentage of frequency in Brittany <sup>18</sup>	Poles in which the mutation is common in Brittany	Possible British and Irish connections <sup>19</sup>
F508del	c. 1000 BC 1,700BC–275 BC	74%	NW-Léon, NE-Trégor, SW-Cap/Bigouden, Aven, Central Vannetais	Wales (71.6%) Scotland (71.3%) Ireland (70.4%)
G551D (the so-called 'Celtic' mutation)	c.800 AD. (900–1550 AD)	3.8%	NW-Léon, NE-Trégor, Aven, Central Vannetais	Wales (2.2%)
1078delT	c. 1000 AD (750–1325 AD)	2.5%	NE-Trégor, SW-Cap/Bigouden, Aven	Wales (2.2%)
1717-IG>A	?	1.7%	NW-Léon, NE-Trégor, SW-Cap/Bigouden, Aven	Wales (1.1 %)
N1303K	?	1.5%	NE-Trégor, SW-Cap/Bigouden, Central Vannetais, SE-Vannetais	No correspondence
G542X	?	1.4%	NW-Léon, NE-Trégor, SW-Cap/Bigouden	Wales (2.2%) Ireland (1%)
1507del	?	1.1%	NE-Trégor	Wales (1.1 %)

17. Future research will examine possible genetic links with regions of England and France.

18. The frequency is given as a percentage in relation to the population of healthy carrier parents in the CFGB database.

19. Bobadilla et al. 2002. Analyses were performed on 72,431 chromosomes and around 100 articles supported the results. As an example, we have noted that the mutations, and their frequency, was mostly encountered in Brittany and across the Channel.

W846X	c. 625 AD 375–1025 AD	1%	NW-Léon, NE- Trégor, Aven, Central Vannetais	No correspond- ence
4005+IG>A	?	0.4%	NW-Léon, NE- Trégor, SW-Cap/ Bigouden, Aven	No correspond- ence
621+IG>T	?	0.3%	Aven, SW-Vann- etais	No correspond- ence
G149R	?	0.2%	NW-Léon, NE- Trégor, SW-Cap/ Bigouden	No correspond- ence

The dates obtained for the most frequent mutations also underline a ‘founder effect’ with a chronological lag necessary to account for the time that the genetic drift occurs within the populations concerned.

Data from Sollicec’s dialectometric analysis, as well as from Pellen’s examination of cystic fibrosis, appear to open a window onto Brittany’s past. Thus, beyond the indirect links on both sides of the Channel, we see a less structured and more diversified migration process taking shape that corresponds more closely to Brett (2011, 2019). The intermingling of populations engendered by this migratory episode created protracted situations of linguistic contact between populations from various regions on the island of Britain and Brittany.

#### 4. Conclusion

The convergence between the results of our studies based on contemporary expression of ancient genetic and linguistic events encouraged us to examine possible historical causes which might explain the parallel phenomena we discovered.

##### 4.1. Analysis of a special case

In this study, our multidisciplinary approach was intended to advance research following the framework of *New Synthesis* school. The results seem to tell the same story from two different perspectives: that of the simultaneous introduction of linguistic varieties of Brittonic and combined genetic input within the Armorican peninsula, especially in the westernmost regions of the peninsula.

This in itself demonstrates the possibility of fruitful exchanges between linguistics and genetic studies.

Population genetics demonstrates that, in the case of genetic diseases, as soon as a gene mutation occurs, its spread or disappearance in a given population is



determined by common social behaviour. In the case of Brittany, the concentration of people sharing the same allele in the same geographical space (in certain cases for millennia) gave birth to genetic pools which were differentiated according to *CFTR* mutations. The genetic pool of a population is thus a reflection of the diversity of its origins but also, and very clearly, the history of the population itself:

Regions with strong cultural particularities tend to separate. This is for example the case for Aquitaine in the South-West which has long represented a civilisation on its own. The Brittany region is also detected as a separate entity in both datasets. This could be explained both by its position at the end of the continent where it forms a peninsula and, by its history since Brittany has been an independent political entity (kingdom and, later, duchy of Bretagne), with stable borders, for a long time.

(Saint Pierre et al. 2020: 862).

In addition, the situation is strongly determined by its demographic dynamics. A small migratory radius, strong homogamy and a population possessing rich genealogical records perfectly illustrates the degree to which the population was structured into a stable nucleus. Historically, the Breton people have in fact experienced several physical and cultural constraints which explain this phenomenon. Firmly anchored in the west of the Armorican peninsula, the consequence of this geographical isolation has resulted in a reduction of genetic diversity by limiting possible external contributions.

Likewise, the application of linguistic distance measurements in the study of vernacular Breton varieties by means of a dialectometric analysis made it possible to observe clear correspondences between levels of linguistic similarity and the distribution of the genetic pools linked to cystic fibrosis. The convergence of our collective findings is most clearly manifested in the radical opposition of the north-western and the south-eastern zones of Breton-speaking Brittany in terms of the linguistic and genetic data.

Our two methods offer a more objective treatment of the data by contrasting the scientific facts and observations which shed light both on the linguistic and genetic present and past of the Breton population. These common observations have thus led us to propose the existence of a 'founder effect' which could well be the consequence of the Brittonic migrations during the early medieval period and, possibly, the absorption of ancient *CFTR* gene mutations carried by the native Armorican population with whom they intermarried.

As we have stated above, this information reinforces the hypotheses of Saint Pierre et al. according to which:

The highest level of SP population in the NW cluster may be either due to its position, at an extreme of the possible Neolithic expansion or to later, and more recent migration from northern Europe with high SP proportion, i.e. Celtic and/or Anglo-Saxon. Exploration of more recent rare genetic variation will allow for a better disentangling of these two hypotheses.

(Saint Pierre et al. 2020: 863)

In the future, we hope to extend the genealogical research based on a highly localised preliminary study of *CFTR* data comparing the Breton and British genetic data. We would then be faced with two types of settlement, both from the NW of Europe by two waves of SP populations: a) a very ancient one dating from the end of the Neolithic and which would be genetically marked by the expansion of the F508del mutation; and b) another migration at the beginning of the Middle Ages which appears to correspond to the spread historical and geographical of the G551D (the so-called ‘Celtic’) and 1078delT (the so-called ‘Welsh’) mutations.

In our view, the scientific approach inherent to genetic and dialectometric research and the concordance of these data appear to not only reinforce many of the hypotheses advanced previously but to open new avenues for future research.

## 5.2. Future perspectives

Multiplying sources is another avenue to explore. More in-depth research could be carried out in south-western Britain on the distribution of cystic fibrosis mutations but also in Normandy where the Brythons are also likely to have migrated. Similarly, in-depth work on the linguistic traits and characteristics of the different poles of linguistic similarity and their possible origin would make it possible to better understand the selection and restructuring process at work in the basins of linguistic traits (Mufwene 2001).

This perspective would allow us to examine how interactions of Brittonic settlers may have occurred with indigenous populations and how these contacts may have triggered the convergence of Gaulish and Brittonic with Late Latin as a supplementary adstratum.

The present work shows the possible benefits of overlapping perspectives of human societies and thus underlines that every individual finds himself at the interface of inherited social practices and biological traits stemming from the distant past but also of a dynamic present that is under unremitting construction. The example of Brittany has allowed us to focus on these phenomena.

*University of Western Brittany (Brest)*

## References

- Barbieri, C., Vicente, M., Oliveira, S., Bostoen, K., Rocha, J., Stoneking, M., & Pakendorf, B.**, 2014, 'Migration and Interaction in a Contact Zone: mtDNA Variation among Bantu-Speakers in Southern Africa', *PLOS ONE, Public Library of Science* 9 (6), available at <https://doi.org/10.1371/journal.pone.0099117> [accessed 14 June 2021].
- Bellis, G., Dehillotte, C., & Lemonnier, L.**, 2019, *Registre Français de la Mucoviscidose, Bilan des données 2017*, Paris: Vaincre la Mucoviscidose — Institut national d'études démographiques.
- Ben Hamed, M., & Darlu, P.**, 2007, 'Gènes et Langues: une longue histoire commune?', *Bulletins et mémoires de la Société d'Anthropologie de Paris* 19 (3–4), 243–264.
- Boas, F.**, 1911, *The Mind of Primitive Man*, New York: The Macmillan Company.
- Bobadilla, J. L., Macek, M. Jr., Fine, J. P., & Farrell, P. M.**, 2002, 'Cystic Fibrosis: a Worldwide Analysis of CFTR Mutations-correlation with Incidence Data and Application to Screening', *Human Mutation* 19, 575–606.
- Bouchard, G., & De Braekeleer, M.**, 1992, *Pourquoi des maladies héréditaires?: Population et génétique au Saguenay-Lac-Saint-Jean*, Sillery: Les éditions du Septentrion.
- Bouget, H., & Coumert, M.**, 2019, 'Enjeux épistémologiques des recherches sur les Breagnes médiévales en histoire, langue et littérature', in: Bouget, H., & Coumert, M., eds., *Histoires des Breagnes 6, Quel Moyen Âge ? La recherche en question*, Brest: CRBC, 15–36.
- Brett, C.**, 2011, 'Soldiers, Saints, and States? The Breton Migrations Revisited', *Cambrian Medieval Celtic Studies* 61, 1–56.
- Brett, C.**, 2019, 'An Invisible Migration? The Origins of Brittany', in: *Migrations et territoires celtiques: Mouvement spatial et mutations culturels / Divroañ hag enbroañ ar Gelted: Dilec'hiañ ha cheñch sevenadur*, Longueil / Rennes: Tir / L'instant même, 11–26.
- Broudic, F.**, 1995, *A la recherche de la frontière. La limite linguistique entre Haute- et Basse-Bretagne aux XIXe et XXe siècles*, Brest: Emgleo Breiz.
- Brun-Trigaud, G.**, 2014, 'Un usage particulier de l'algorithme de Damerau-Levenshtein dans le domaine occitan', in: Diemoz, F., Aquino-Weber, D., Grüner L., & Reusser-Elzingre, A., eds., *"Toujours langue varie..." Mélanges de linguistique historique du français et de dialectologie galloromane offerts à M. le Professeur Andres Kristol par ses collègues et anciens élèves* (Recueil de Travaux publiés par la Faculté des Lettres et Sciences humaines de l'Université de Neuchâtel 59), Genève: Droz, 127–148 (available at <https://halshs.archives-ouvertes.fr/halshs-01067295/document> [last accessed 26 May 2018]).
- Brun-Trigaud, G., Sollic, T., & Le Dù, J.**, 2016, 'A New Dialectometric Approach Applied to the Breton Language', in: Côté, M.-H., Knooihuizen, R., & Nerbonne, J., *The Future of Dialects: Selected Papers from Methods in Dialectology 15. Language Variation*, 135–54, Berlin: Language Science Press (available at 10.17169/langsci.b81.147 [last accessed 16 March 2016]).

- Calafell, de, F.**, et al., 2004, 'The Molecular Genetic Epidemiology of Cystic Fibrosis Report of a Joint Meeting of WHO/ECFTN/ICF(M)A/ECFS', Genoa, Italy, 19 June 2002 (available at [https://www.cfww.org/docs/who/2002/who\\_hgn\\_cf\\_wg\\_04.02.pdf](https://www.cfww.org/docs/who/2002/who_hgn_cf_wg_04.02.pdf) [last accessed 25 September 2020]).
- Campbell, L.**, 2015, 'Do Languages and Genes Correlate?', *Language Dynamics and Change* 5 (2), 202–226 (available at <https://doi.org/10.1163/22105832-00502007> [last accessed 14 June 2021]).
- Cassard, J.-C.**, 2003, *Les Bretons de Nominoë*, Rennes: Presses Universitaires de Rennes.
- Cavalli-Sforza, L., Piazza, A., Menozzi, P., & Mountain, J.**, 1989, 'Genetic and Linguistic Evolution', *Science. American Association for the Advancement of Science* 244 (4909), 1128–1129 (available at <https://doi.org/10.1126/science.2727697> [last accessed 14 June 2021]).
- Cavalli-Sforza, L.**, 1996, *Gènes, peuples et langues*, Paris: Odile Jacob.
- Chadwick, N.**, 1969, *Early Brittany*, Cardiff: University of Wales Press.
- Cheadle, J.P., Goodchild, M.C., & Meredith, A.L.**, 1993, 'Direct Sequencing of the Complete CFTR Gene: the Molecular Characterisation of 99.5% of CF Chromosomes in Wales', *Human Molecular Genetics* 2, 1551–1556.
- Chédeville, A., & Guillotel, H.**, 1984, *La Bretagne des Saints et des Rois, Ve-Xe siècle*, Rennes: Ouest-France.
- Costaouec, D.**, 1998, *Le breton parlé à la forêt-Fouesnant (Finistère-sud) : pratique actuelle. Problèmes de phonologie et de syntaxe*, Paris 5 — René Descartes thèse de doctorat nouveau régime.
- Costaouec, D.**, 2012, 'Linguistic Geography of Breton and Sociocultural Motivations', *STUF—Language Typology and Universals* 65 (1), 47–64.
- Coumert, M.**, 2010, 'Le peuplement de l'Armorique : Cornouaille et Domnonée de part et d'autre de la Manche aux premiers siècles du Moyen Age', in: Coumert, M., & Tétrel, H., eds., *Histoires des Breagnes 1. Les mythes fondateurs*, Brest: CRBC, 15–22.
- Estivill, X., Bancells, C., & Ramos, C.**, 1997, 'Geographic Distribution and Regional Origin of 272 Cystic Fibrosis Mutations in European Populations', *Human Mutation* 10 (2), 135–54 (available at [https://doi.org/10.1002/\(SICI\)1098-1004\(1997\)10:2<135::AID-HUMU6>3.0.CO;2-J](https://doi.org/10.1002/(SICI)1098-1004(1997)10:2<135::AID-HUMU6>3.0.CO;2-J) [last accessed 25 September 2020]).
- Evans, D. E.**, 1990, 'Insular Celtic and the Emergence of the Welsh Language', in: Bammesberger, A., & Wollmann, A., eds., *Britain 400–600: Language and History*, Heidelberg: Carl Winter, 149–178.
- Excoffier, L., Langaney, A., Dupanloup, I., Polom, E. S., & Schneider, S.**, 2002, 'Génétique, linguistique et histoire des peuplements humains', *Languages* 36 (146), 80–90 (available at <https://doi.org/10.3406/lgge.2002.2403> [last accessed 25 September 2020]).
- Falc'hun, F.**, 1981, *Perspectives nouvelles sur l'histoire de la langue bretonne* (La Nation En Question), Paris: Union Générale d'Éditions.

- Falileyev**, A., 2003, 'Languages of Old Wales: A Case for Co-existence', *Dialectologia et Geolinguistica* 11, 18–38 (available at <https://doi.org/10.1515/dig.2003.2003.11.18> [last accessed 25 September 2020]).
- Farrell**, P., **Férec**, C., **Macek**, M., **Frischer**, T., **Renner**, S., **Riss**, K., **Barton**, D., et al., 2018, 'Estimating the Age of p.(Phe508del) with Family Studies of Geographically Distinct European Populations and the Early Spread of Cystic Fibrosis', *European Journal of Human Genetics, Nature Publishing Group* 26 (12), 1832–1839 (available at <https://doi.org/10.1038/s41431-018-0234-z> [last accessed 14 June 2021]).
- Fichou**, Y., **Génin**, E., **Le Maréchal**, C., **Audrézet**, M. P., **Scotet**, V., & **Férec**, C., 2008, 'Estimating the Age of *CFTR* Mutations Predominantly Found in Brittany (Western France)', *Journal of Cystic Fibrosis* 7 (2), 168–173 (available at <https://doi.org/10.1016/j.jcf.2007.07.009> [last accessed 14 June 2021]).
- Fleuriot**, L., 1980, *Les origines de la Bretagne. L'émigration*, Paris: Payot.
- German**, G., 1984, *Une étude linguistique sur le breton de Saint-Yvi*, Brest: Université de Bretagne Occidentale. Doctorat de 3e cycle.
- German**, G., 1993, Une méthode dialectométrique (assistée par ordinateur) pour l'analyse des atlas linguistiques, *La Bretagne linguistique* 7, 177–213.
- German**, G., 2006, 'Anthroponyms as Markers of "Celticity" in Brittany, Cornwall and Wales', in: Tristram, H. L. C., ed., *The Celtic Englishes IV. The Interface between English and the Celtic Languages. Proceedings of the 4<sup>th</sup> International Colloquium on the "Celtic Englishes" held at the University of Potsdam in Golm (Germany) from 22–26 September 2004*, Potsdam: Universitätsverlag Potsdam, 34–63.
- German**, G., 2008, 'Language Shift, Diglossia and Dialectal Variation in Western Brittany: The Case of Southern Cornouaille', in: Tristram, H. L. C., ed., *The Celtic Languages in Contact: Papers from the workshop within the framework of the XIII International Congress of Celtic Studies, Bonn, July 2007*, Potsdam: Universitätsverlag Potsdam, 146–192.
- German**, G., 2014, 'Breton Patronyms and the British Heroic Age', in: Hellegouarc'h, A., & Williams, H., eds., *Regards croisés sur la Bretagne et le pays de Galles/Cross-cultural Essays on Wales and Brittany*, Brest: Centre de Recherche Bretonne et Celtique; Aberystwyth: Centre for Advanced Celtic and Welsh Studies, 53–88.
- Gildas**, 1978, *The Ruin of Britain and other documents* (History from the Sources, trans. Winterbottom, M.) Guildford, Surrey: Phillimore.
- Goebel**, H., 1982, *Dialektometrie: Prinzipien und Methoden des Einsatzes der Numerischen Taxonomie im Bereich der Dialektgeographie*, Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Goebel**, H., 1983, 'Parquet polygonal et treillis triangulaire, les deux versants de la dialectométrie interponctuelle', *Revue de linguistique romane* 47, 353–412.
- Goebel**, H., 1984, *Dialektometrische Studien: Anhand italoromanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF*, Tübingen: Max Niemeyer Verlag.

- Goebel**, H., 2010, 'Dialectometry and Quantitative Mapping', in: Lameli, A., Kehrein R., & Rababus, S., eds., *Language and Space. An International Handbook of Linguistic Variation* (Vol. 2, *Language Mapping*), Berlin, Boston: De Gruyter Mouton, 433–457.
- Guy**, B., 2014, 'The Breton Migration: a New Synthesis', *Zeitschrift für celtische Philologie* 61, 101–156.
- Heeringa**, W., 2004, *Measuring Dialect Pronunciation Differences using Levenshtein Distance*, unpublished Ph.D. dissertation, Groningen: University of Groningen (available at <http://citeseerx.ist.psu.edu/showciting.jsessionid=-328FA9F7293B9025236DF0AAEA070600?cid=36229003> [last accessed 15 March 2016]).
- Humphreys**, H., 1991, 'The Geolinguistics of Breton', in: Williams, C. H., ed., *Linguistic Minorities, Society and Territory*, Clevedon: Multilingual Matters, 96–119.
- Jackson**, K. H., 1967, *A Historical Phonology of Breton*, Dublin: DIAS.
- Labov**, W., 1974, 'On the Use of the Present to Explain the Past', in: *Proceedings of the Eleventh International Congress of Linguists*, Bologna: Il Mulino, 825–851.
- Labov**, W., 1994, *Principles of Linguistic Change. Vol. 1. Internal Factors, Language*, Oxford: Blackwell Publishing.
- Le Berre**, Y., & **Le Dù**, J., 1996, 'Parité et disparité : sphère publique et sphère privée de la parole', *La Bretagne linguistique* 10, *Actes du colloque Badume-Standard-Norme, Le double jeu de la langue*, Brest: CRBC, 7–25.
- Le Dù**, J., 1972, 'Le Nouvel Atlas Linguistique de Basse-Bretagne', *Etudes celtiques* 13 (1), 332–345.
- Le Dù**, J., 1988, 'Les appellatifs ethniques dans l'anthroponymie de la Basse-Bretagne, *Mélanges offerts à Pierre Flatrès*, Lille: Sociétés des Géographes de Lilles, 67–76.
- Le Dù**, J., 2001, *Nouvel Atlas Linguistique de la Basse-Bretagne*, 2 vols., Brest: CRBC.
- Le Roux**, P., 1924, *Atlas linguistique de la Basse-Bretagne*, Rennes: Plihon et Hommay.
- Levenshtein**, V. I., 1966, 'Binary Codes Capable of Correcting Deletions, Insertions and Reversals', *Soviet Physics. Doklady* 10, 707.
- Longobardi**, G., **Ghirotto**, S., **Guardiano**, C., **Tassi**, F., **Benazzo**, A., **Ceolin**, A., & **Barbujani**, G., 2015, 'Across Language Families: Genome Diversity Mirrors Linguistic Variation within Europe', *American Journal of Physical Anthropology* 157 (4), 630–640 (available at <https://doi.org/10.1002/ajpa.22758> [last accessed 14 June 2021]).
- Loth**, J., 1883, *L'émigration bretonne en Armorique, du Ve au VIIe siècle de notre ère*, Rennes: E. Baraise et cie.
- Manni**, F., & **Barrai**, I., 2001, 'Genetic Structures and Linguistic Boundaries in Italy: A Microregional Approach', *Human Biology* 73(3), 335–347.
- Manni**, F., **Heeringa**, W., & **Nerbonne**, J., 2006, 'To What Extent are Surnames Words? Comparing Geographic Patterns of Surname and Dialect Variation in the Netherlands', *Literary and Linguistic Computing* 21(4), 507–527 (available at <https://doi.org/10.1093/lc/fql040> [last accessed 14 June 2021]).



- Mayr, E.**, 1942, *Systematics and the Origin of Species, from the Viewpoint of a Zoologist*, Cambridge, MA: Harvard University Press.
- Mayr, E.**, 1954, 'Change of Genetic Environment and Evolution', in: Huxley, J., ed., *Evolution as a Process*, London: George Allen & Unwin, 157–180.
- Mayr, E.**, 1982, *The Growth of Biological Thought. Diversity, Evolution, and Inheritance*, Belknap Press of Harvard University Press.
- McMahon, R.**, 2004, 'Genes and Languages', *Public Health Genomics* 7(1), 2–13 (available at <https://doi.org/10.1159/000080298> [last accessed 14 June 2021]).
- Métoz, L.**, 2006, 'L'émergence de la lexicostatistique et de la glottochronologie en France au XIXe siècle: Constantine Samuel Rafinesque, Jules Dumont d'Urville et Paul Broca', *Histoire Épistémologie Langage* 28(2), 125–155 (available at <https://doi.org/10.3406/hel.2006.2886> [last accessed 14 June 2021]).
- Mufwene, S. S.**, 2001, *The Ecology of Language Evolution*, Cambridge: Cambridge University Press.
- Nerbonne, J.**, & **Heeringa, W.**, 1997, 'Measuring Dialect Distance Phonetically', in: Coleman, J., ed., *Workshop on Computational Phonology*, Madrid: Association for Computational Linguistics, 11–18 (available at <http://aclweb.org/anthology/W/W97/W97-1102>, 10 May, 2018 [last accessed 14 June 2021]).
- Nerbonne, J.**, & **Heeringa, W.**, 2010, 'Measuring Dialect Differences', in: Auer, P., & Schmidt, J., eds., *Language and Space. An International Handbook of Language Variation — Vol. 1 Theories and Methods*, 550–567, Berlin: De Gruyter Mouton.
- Nerbonne, J.**, **Heeringa, W.**, & **Kleiweg, P.**, 1999, 'Comparison and Classification of Dialects', in: *Proceedings of the Ninth Conference on European Chapter of the Association for Computational Linguistics*, 281–282.
- Nettle, D.**, & **Harriss, L.**, 2003, 'Genetic and Linguistic Affinities between Human Populations in Eurasia and West Africa', *Human Biology*, Wayne State University Press 75(3), 331–344 (available at <https://doi.org/10.1353/hub.2003.0048> [last accessed 14 June 2021]).
- Pakendorf, B.**, 2014, 'Coevolution of Languages and Genes', *Current Opinion in Genetics & Development* (Genetics of Human Evolution) 29, 39–44 (available at <https://doi.org/10.1016/j.gde.2014.07.006> [last accessed 14 June 2021]).
- Pellen, N.**, 2015, *La mucoviscidose en héritage*, Paris: INED.
- Provine, W. B.**, 2004, 'Ernst Mayr: Genetics and Speciation', *Genetics* 167 (3), 1041–1046.
- Riordan, J. R.**, **Rommens, J. M.**, **Kerem, B.**, **Alon, N.**, **Rozmahel, R.**, **Grzelczak, Z.**, **Zielenski, J.**, et al., 1989, 'Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA', *Science*, American Association for the Advancement of Science 245 (4922), 1066–1073 (available at <https://doi.org/10.1126/science.2475911> [last accessed 14 June 2021]).

- Rommens, J. M., Iannuzzi, M. C., Kerem, B., Drumm, M. L., Melmer, G., Dean, M., Rozmahel, R., et al.**, 1989, 'Identification of the Cystic Fibrosis Gene: Chromosome Walking and Jumping', *Science*, American Association for the Advancement of Science 245 (4922), 1059–1065 (available at <https://doi.org/10.1126/science.2772657> [last accessed 14 June 2021]).
- Saint Pierre, A., Giemza, J., Alves, I., Karakachoff, M., Gaudin, M., Amouyel, P., Dartigues, J.-F., et al.**, 2020, 'The Genetic History of France', *European Journal of Human Genetics* 28 (7), 853–865 (available at <https://doi.org/10.1038/s41431-020-0584-1> [last accessed 14 June 2021]).
- Schreier, D.**, 2017, 'Dialect Formation in Isolated Communities', *Annual Review of Linguistics* 3 (1), 347–362 (available at <https://doi.org/10.1146/annurev-linguistics-011516-034115> [last accessed 14 June 2021]).
- Scotet, V., Barton, D. E., Watson, J. B. G., Audrézet, M.P., McDevitt, T., McQuaid, S., Shortt, C., De Braekeleer, M., Férec C., & Le Maréchal, C.**, 2003, 'Comparison of the *CFTR* Mutation Spectrum in Three Cohorts of Patients of Celtic Origin from Brittany (France) and Ireland', *Human Mutation* 22 (1), 105 (available at <https://doi.org/10.1002/humu.9158> [last accessed 14 June 2021]).
- Segalen, M.**, 1985, *Quinze générations de Bas-Bretons. Parenté et société dans le pays bigouden Sud, 1720-1980*, Paris: Presses universitaires de France.
- Séguy, J.**, 1971, 'La relation entre la distance spatiale et la distance lexicale', *Revue de linguistique romane* 35 (139–140), 335–357.
- Séguy, J.**, 1973a, 'La dialectométrie dans l'Atlas linguistique de la Gascogne', *Revue de linguistique romane* 37 (145–146), 1–24.
- Séguy, J.**, 1973b, *Atlas linguistique de la Gascogne*, vol. VI, Paris: Editions du Centre national de la recherche scientifique.
- Serva, M.**, 2012, *The Settlement of Madagascar: What Dialects and Languages Can Tell Us*. *PLoS ONE* 7 (2). e30666 (available at <https://doi.org/10.1371/journal.pone.0030666> [last accessed 14 June 2021]).
- Serva, M., Petroni, F., Volchenkov, D., & Wichmann, S.**, 2012, 'Malagasy Dialects and the Peopling of Madagascar', *Journal of The Royal Society Interface* 9 (66), 54–67 (available at <https://doi.org/10.1098/rsif.2011.0228> [last accessed 14 June 2021]).
- Sherriah, A. C., Devonish, H., Thomas, E. A. C., & Creanza, N.**, 2018, 'Using Features of a Creole Language to Reconstruct Population History and Cultural Evolution: Tracing the English Origins of Sranan', *Philosophical Transactions of the Royal Society B: Biological Sciences* 373 (1743) (available at <https://doi.org/10.1098/rstb.2017.0055> [last accessed 14 June 2021]).
- Sims-Williams, P.**, 1998, 'Genetics, Linguistics, and Prehistory: Thinking Big and Thinking Straight', *Antiquity* 72 (277), 505–527 (available at <https://doi.org/10.1017/S0003598X00086932> [last accessed 14 June 2021]).



- Sims-Williams, P.**, 2012, 'Bronze- and Iron-Age Celtic-speakers: What Don't We Know, What Can't We Know, and What Could We Know? Language, Genetics and Archaeology in the Twenty-first Century', *The Antiquaries Journal* 92, 427–49 (available at <https://doi.org/10.1017/S000358151200011X> [last accessed 14 June 2021]).
- Slon, V., Mafessoni, F., Vernot, B., de Filippo, C., Grote, S., Viola, B., Hajdinjak, M.**, et al., 2018, 'The Genome of the Offspring of a Neanderthal Mother and a Denisovan Father', *Nature* 561 (7721), 113–116 (available at <https://doi.org/10.1038/s41586-018-0455-x> [last accessed 14 June 2021]).
- Solliec, T.**, 2017, 'Explorer la variation en breton grâce à la dialectométrie : la Basse-Bretagne considérée par la distance linguistique', *La Bretagne linguistique* 21, 29–61.
- Solliec, T.**, 2019, 'Unity in Diversity? A Fine-grained Approach to the Linguistic Geography of Breton by the Means of Dialectometry', *Studia Celto-Slavica* 10, 137–173.
- Solliec, T.**, 2021, *Distance linguistique et dialectométrie, une application à la langue bretonne. Enjeux, méthodologie et interprétations*, unpublished PhD dissertation, Brest: University of Western Brittany.
- Tristram, H. L. C.**, 1993, 'Zwiebeln und Wörter: Zum Sprachkontakt über den Ärmelkanal', in: Zimmer, S., & Rockel, M., eds., *Akten des ersten Symposiums deutschsprachiger Keltologen in Berlin, 8.-10. April 1992*, Tübingen: Niemeyer, 13–51.
- Tristram, H. L. C.**, 1995a, 'Linguistic Contacts across the English Channel: The Case of the Breton Retroflex <r>', in: Fisiak, J., ed., *Linguistic Change under Contact Conditions*, Berlin: de Gruyter, 291–313.
- Tristram, H. L. C.**, 1995b, 'Zaoz and Zomerzet: Linguistic Contacts Across the English Channel', in: Viereck, W., ed., *Verhandlungen des Internationalen Dialektologenkongresses Bamberg 1990 (Zeitschrift für Dialektologie und Linguistik 77)*, Stuttgart: Steiner, 276–289.
- Trudgill, P.**, 2004, *New-dialect Formation. The Inevitability of Colonial Englishes*, Edinburgh: Edinburgh University Press.
- UK Cystic Fibrosis Registry**, 2019, *2018 Annual Data Report* (available at <https://www.cysticfibrosis.org.uk/~media/documents/the-work-we-do/uk-cf-registry/2018-registry-annual-data-report.ashx?la=en> [last accessed 22 August 2020]).
- Zolin, A., Orenti, A., Naehrlich L., van Rens, J.**, et al., 2019, *ECFSR Annual Report, 2017*.

Annex 1

Location of couples with the maximum number of carriers by mutation among their descendants

