

The Genetic Structure of a Highland Clan

Bryan Sykes and Jayne Nicholson

University of Oxford
Weatherall Institute of Molecular Medicine
Oxford OX3 9DS

Keywords: Y-chromosome, surnames, Scottish clans

Co-inheritance of Y-chromosomes and surnames can be used to test the historical accuracy of traditional genealogies and to reconstruct the haplotype of a common ancestor. Here we test the genealogy of the most influential of the Scottish Highland clans, Clan Donald, and show that, contrary to expectations, large numbers of present day members of the clan are directly descended from the 12th century Celtic hero *Somhairle mac Gillebride* (Sommerled) whose descendants founded the clan.

Where the inheritance of surnames is patrilineal there should, in theory, be an association between the name and the non-recombining segment of the Y-chromosome, which is also passed from father to son. In the descendants of a specific ancestor, the association will be disrupted by non-paternity events, such as adoption or adultery, and it was thought that these disrupting influences would combine to erase a strong connection between surnames and Y-chromosomes in present day populations ¹. However, it was later shown that over half of a randomly ascertained sample of English men with the same surname (Sykes) shared an identical, or very similar, Y-chromosome, indicative of a single major founder ². Like many English names, it became hereditary in the 13th century and the averaged rate of non-paternity events needed to retain the observed association over eight centuries is, at 1.3% per generation, much lower than presently. Many other English surnames have since been shown to have a similarly high, or even higher, association with particular Y-chromosomes ³. Scottish clan names, however, have very different origins. Although also following a patrilineal inheritance, name adoption was almost universal in the Highlands as men took the name of the clan chief on whose lands they lived or with whom they fought. This predicts a wide variety of Y-chromosomes among present day members, mostly unrelated to the founder of the clan. To investigate this proportion and to show how a clan's genetic structure compares to its traditional genealogy, we chose Clan Donald, the largest of the Highland clans with an estimated membership worldwide of 3-4 million. Here we show that a substantial proportion of these are direct patrilineal descendants of the clan founder himself.

Clan Donald takes its name from Donald of Islay from whom all branches of the clan, both extant and extinct, trace their direct patrilineal descent (Figure 1). According to clan histories, Donald of Islay was himself the grandson of the 12th century hero *Somhairle mac Gillebride*, known as Sommerled, and his second wife Ragnhilde, daughter of the King of Man. Sommerled is credited with regaining Gaelic political control of the west coast and the islands from the Norse kings. Another clan, the MacDougalls of Lorn, also claim descent from Sommerled through his son Dugall (Figure). The political and cultural influence of the two clans on the western seaboard of Scotland was immense and culminated in their tenure of the autonomous Lordship of the Isles. Following the defeat of the Charles Edward Stewart at Culloden in 1746, the power of Clan Donald and other Highland clans declined and many clansmen emigrated overseas. The exodus was continued by economic depression and by the Highland Clearances. As clan numbers declined in Scotland, they greatly increased overall as the emigrants prospered.

Y-chromosomes from 89 males with the surnames Macdonald(64), MacDougall(13) and Macallister(12), and living in Argyll, Highland, Skye or the Western Isles, collected either at blood donor sessions or in response to a mailed request, were typed at a range of

biallelic and microsatellite markers. As a control, 205 Y-chromosomes from the same region but from without the clan were similarly analysed. Four clusters of inter-related Y-chromosomes were identified, clusters being defined as embracing all haplotypes within the same haplogroup which were either identical to, or one mutational step removed from, other members of the same cluster. The largest, cluster A, comprised 48 chromosomes from haplogroup 1 with 23 different haplotypes varying in frequency from 1-7. Cluster B comprised 22 chromosomes from haplogroup 3 with four different haplotypes whose frequencies ranged from 1 to 17 occurrences. Cluster C comprised four haplogroup 3 chromosomes containing two different haplotypes with frequencies of 1 and 3 occurrences. Cluster D contained only two chromosomes. The remaining 13 chromosomes were single occurrences at least two mutations distant from any other.

Chromosomes in cluster A are associated with all three surnames but there is no distinct modal haplotype indicative of a common origin for the cluster. Cluster C chromosomes are only carried by Macdonalds. Cluster D comprises only two individuals, both Macdougalls. In contrast, Cluster B chromosomes had two features compatible with a descent from a common ancestor, probably Somerled. Firstly, the cluster comprises a frequent modal haplotype accommodating 17 of the 22 chromosomes with the remaining chromosomes differing at only one marker (Table 1). Secondly, the modal haplotype is associated with all three surnames. Thirdly, outside these names it is rare in this region of Scotland with only one occurrence in the 205 controls, and that in a man with the surname MacArthur. The MacArthurs of Islay, hereditary pipers to the Clan Donald, also claim descent from Somerled. The substantially increased frequency of the modal and closely related haplotypes, not only within Clan Donald but also among the Macdougalls, is a strong piece of circumstantial evidence that cluster B chromosomes may have been inherited from Somerled himself.

To confirm this speculation, DNA samples were sought, and received, from the clan chiefs of the five extant branches of Clan Donald whose genealogies claim direct patrilineal descent from Somerled. These are Lord Macdonald, Sir Ian McDonald of Sleat, Ranald Macdonald of Clanranald, Ranald MacDonell of Glengarry and William McAlester of Loup. Their Y-chromosomes were typed for the same microsatellite haplotypes and the results are shown in Table 2. Four chromosomes were identical not only to each other but to the modal haplotype identified in Cluster B. The fifth chromosome has a difference at one microsatellite locus compatible with an identical origin 900 years ago. This leaves little room for doubt that we have identified the Y-chromosome of Somerled among his direct ancestors.

Several interesting points flow from this conclusion. First is the remarkable consistency in the genealogies of the five chiefs. There are a total of 85 independent father-son transmissions in the Clan Donald genealogy without a single observed non-paternity event. Second is the abundance of Somerled's Y-chromosome among other members of the Clan Donald, and the Macdougalls. In this study 19% of men within the two clans claiming descent from Somerled have inherited his Y-chromosome with a further 6% having a very closely related chromosome. If the sample is representative and the estimate of the present day membership of the clan is accurate, there may be as many as a quarter of a million direct descendants of Somerled who carry his Y-chromosome.

What can we tell of Somerled's own patrilineal descent? Traditional genealogies trace Somerled's paternal ancestry through his father Gillebride to a long line of Irish kings back to the fourth century Colla Uais, founder of the kingdom of Oriel and, ultimately, to the legendary Conn of a Hundred Battles who lived in the 2nd century AD ⁷. However, the haplotype of Somerled's Y-chromosome deduced from this study, points to a quite different paternal ancestry. Haplogroup 3, to which Somerled's chromosome belongs, is extremely rare in Ireland, where haplogroup 1 chromosomes are almost fixed among those with Gaelic surnames ⁸. However, haplogroup 3 chromosomes, are widespread in Norway and the former Norse colony of Shetland. The Somerled modal haplotype itself is found at a frequency of 3% in a sample of Norway and 1.5% in Shetland. On this evidence Somerled, the Celtic hero who swept the Norse from the Isles, was himself descended, not from a Celt, but from a Viking.

* For convenience all variations (McDonald, McDougall, McAllister etc.,) are unified under the prefix Mac.

Acknowledgements

We are very grateful to all the volunteers who participated in this study and to the staff and directors of the Scottish Blood Transfusion Service for their assistance. We also want to thank Margaret Macdonald and Anne Mackinnon from the Clan Donald Library at Armadale, Isle of Skye for their advice. This work was supported by grants from the Wellcome Trust and the Medical Research Council and Oxford Ancestors Limited.

Competing Interest Statement

BCS is chairman of Oxford Ancestors Limited.

Correspondence should be addressed to Bryan Sykes (bryan.sykes@wolfson.ox.ac.uk)

.

References

- 1 Jobling M, Pandya A, Tyler-Smith C (1997) The Y-chromosome in forensic analysis and paternity testing. *Int.J.Legal.Med* **110**, 118-124.
- 2 Sykes BC, Irven CI (2000) Surnames and the Y-chromosome. *Am.J.Hum.Genet.* **66**, 1417-1419
- 3 Sykes BC. Unpublished data.
- 4 Rosser Zh, Zerjal T, Hurles ME et al. (2000) Y-chromosomal diversity in Europe is clinal and influenced primarily by geography rather than language. *Am.J.Hum.Genet* **67**, 1526-1543
- 5 de Knieff P et al (1997) Chromosome Y microsatellites: population genetics and evolutionary aspects. *Int.J.Legal.Med* **110**: 134-140.
- 6 Sykes BC. Unpublished data.
- 7 Marsden J. (2000). *Somerled and the emergence of Gaelic Scotland*. Tuckwell Press, East Linton, Midlothian.
- 8 Hill EW et al. (2000) Y-chromosomes and Irish origins. *Nature* **404**, 351-352.

Table 1 **Clan Donald Y-chromosomes from Cluster B**

Surname	N	19	389-I	389-II	390	391	392	393
Macdonald	7	15	11	17	25	11	11	13
Macdougall	4	15	11	17	25	11	11	13
Macallister	6	15	11	17	25	11	11	13
Macdonald	1	15	10	17	25	11	11	13
Macdonald	1	15	11	17	25	10	11	13
Macdonald	3	15	11	17	27	11	11	13

Table 2 **Y-chromosomes of Clan Donald chiefs**

Name	19	389-I	389-II	390	391	392	393
Lord Macdonald	15	11	17	25	11	11	13
Sir Ian Macdonald of Sleat	15	11	17	25	11	11	13
Ranald Macdonald of Clanranald	15	11	17	25	11	11	13
Ranald MacDonell of Glengarry	15	11	17	23	11	11	13
Willaim McAlester of Loup	15	11	17	25	11	11	13

Microsatellite repeat lengths. All loci are prefixed by DYS. Alleles which differ from the modal haplotype are shown in bold.

Figure Legend

Clan Donald genealogy showing the main branches leading to the present clan chiefs. Numbers of father-son transmissions connecting chiefs to the nodes are shown in bold. The total of father-son transmissions in this genealogy is 87.

