

**Supplementary Table 2a.** Number of populations in grid of ethnicity, linguistic lineage and geographical zone

		Social Category	Linguistic Category				Total	Total by zone
			IE	DR	AA	TB		
ZONE	N	Large	15	-	-	2	17	
	N	Isolated	2	-	-	1	3	20
	S	Large	0	5	-	-	5	
	S	Isolated	1	4	-	-	5	10
	E	Large	3	-	-	-	3	
	E	Isolated	1	1	3	-	5	8
	W	Large	4	-	-	-	4	
	W	Isolated	2	-	1	-	3	
	W	Out-group	1	-	-	-	1	8
	C	Large	-	-	-	-	0	
	C	Isolated	-	2	3	-	5	5
	NE	Large	1	-	-	1	2	
	NE	Isolated	1	-	1		2	4
		Total		31	12	8	4	55

**Supplementary Table 2b.** Total number of samples in each grid of ethnicity, linguistic lineage and geographical zone

		<b>Social Category</b>	<b>IE</b>	<b>DR</b>	<b>AA</b>	<b>TB</b>	<b>Total</b>	<b>Total by zone</b>	<b>Total</b>	
ZONE	N	Large	611					611		
	N	Isolated	87			119		206	817	
	S	Large		184				184		
	S	Isolated	46	78				124	308	
	E	Large	105					105		
	E	Isolated	28	39	62			129	234	
	W	Large	126					126		
	W	Isolated	44		19			63		
	W	Out-group	23					23	212	
	C	Large						0		
	C	Isolated		68	87			155	155	
	NE	Large	22			40		62		
	NE	Isolated	42		41			83	145	
	Total			1134	369	209	159		1871	1871

The distribution of samples which have been considered for final validation