

Some basic statistics related to transliteration files using STA codes.

These are valid for Intermediate STA (15), 16/04/2020

	CD	FG	IT	GC	ZL
Count	2196	4060	5215	5367	5389

Table 1: counts of the number of loci covered by the main transliteration files

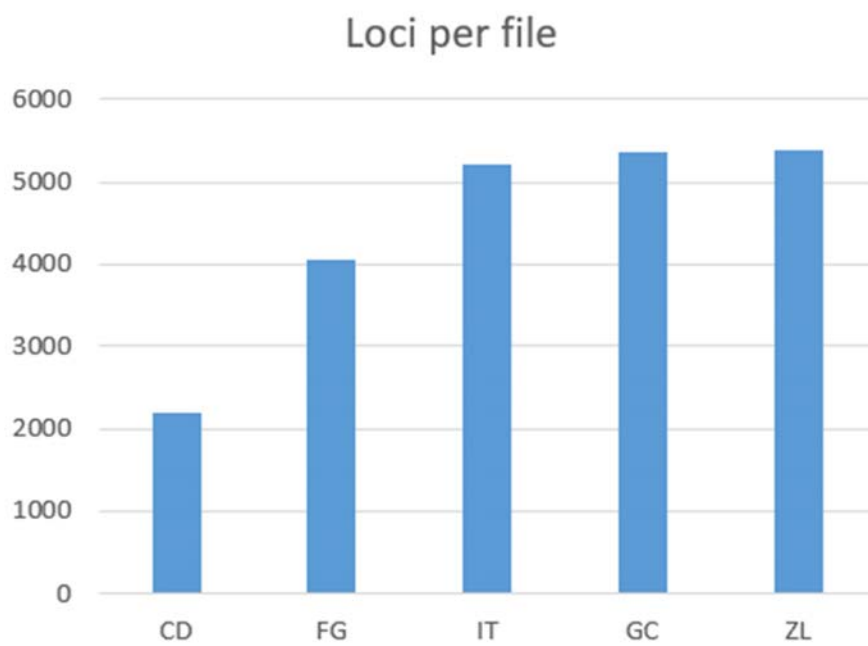


Figure 1: counts for the number of loci covered by the main transliteration files

	CD	FG	IT	GC	ZL
Count	65,162	135,532	155,272	157,089	157,202

Table 2: counts of the number of STA characters included in the main transliteration files

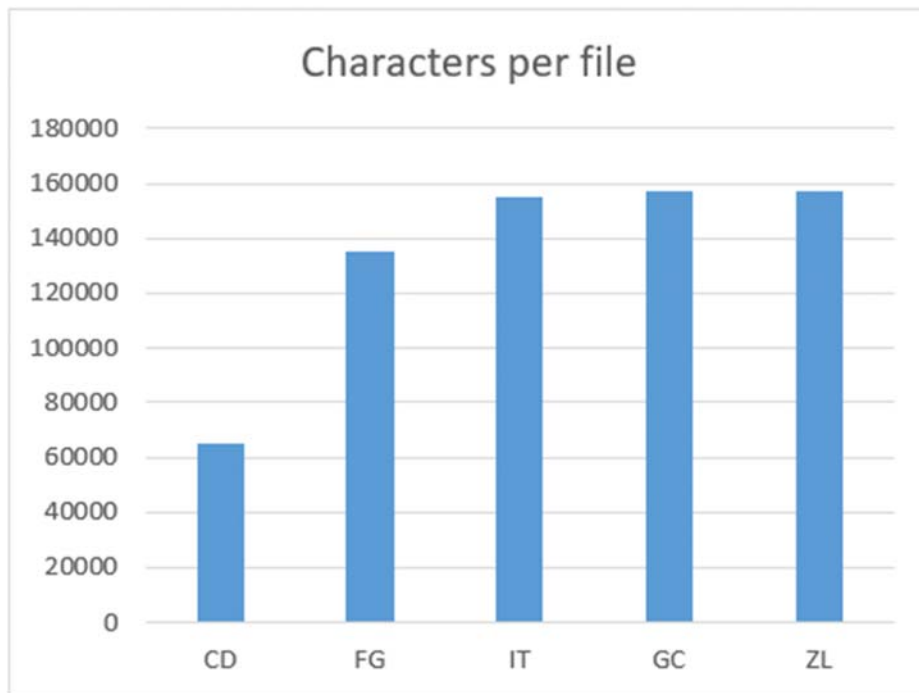


Figure 2: counts of the number of STA characters included in the main transliteration files

Family	CD	FG	IT	GC	ZL
A	23724	49548	57378	58199	58323
B	10815	21677	24705	25107	25144
C	3693	7904	9601	9608	9669
D	659	1579	1844	1885	1876
E	6481	13711	16020	16150	16132
F	4099	8917	10430	10362	10237
G	6447	13648	15363	15498	15536
H	2270	4174	4502	4680	4590
J	98	280	394	432	423
K	4	3	14	18	15
L	2563	5187	5423	5427	5432
M	140	251	395	325	342
N	916	1029	2440	2461	2424
P	1774	5142	4235	4293	4322
Q	89	40	157	186	169
R	155	293	271	290	325
S	0	0	20	26	32
U	1066	1887	1804	1879	1872
V	0	0	54	89	120
X	0	25	44	84	74
Z	169	237	178	90	145
Total	65,162	135,532	155,272	157,089	157,202

Table 3: counts of the STA codes per family for the main transliteration files

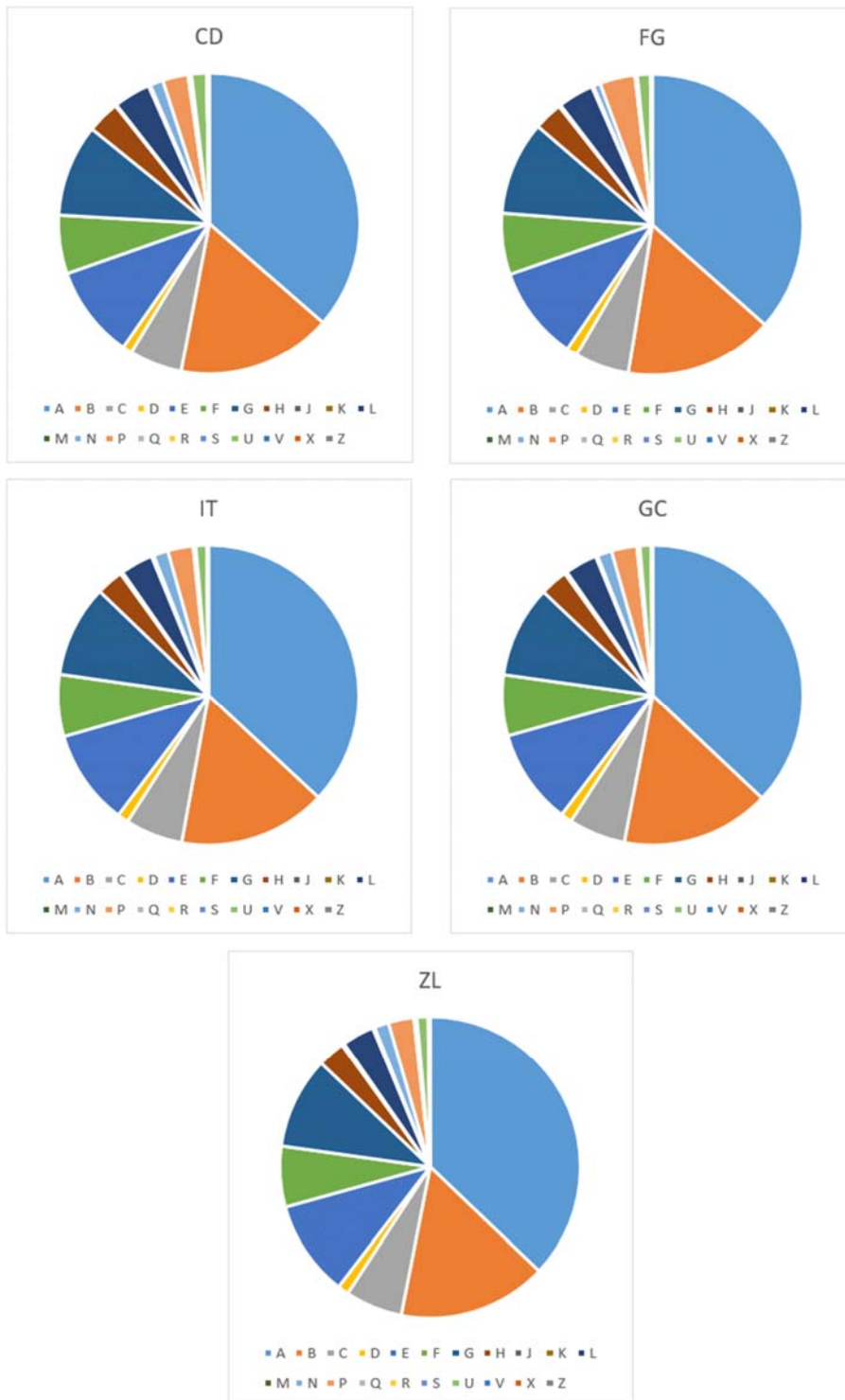


Figure 3: distribution of STA family codes over the main transliteration files, based on Table 3

	CD	FG	IT	GC	ZL
CD		96.9	98.2	98.2	98.4
FG	97.2		97.8	97.6	97.8
IT	97.8	96.8		98.5	98.5
GC	98.0	96.8	98.8		99.0
ZL	98.2	97.0	98.8	99.0	

Table 4: similarity in terms of STA families between two transliterations, expressed in % of the row (label).
 Example: 97.0% of all characters in the ZL transliteration have the same family code as those in the FG transliteration (light green cell).

	CD	FG	IT	GC	ZL
CD		95.4	97.0	91.4	97.4
FG	95.7		96.8	90.5	96.5
IT	96.6	95.8		91.4	97.3
GC	91.2	89.7	91.7		92.0
ZL	97.2	95.7	97.6	91.9	

Table 5: similarity in terms of STA codes between two transliterations (regularisation level 1), expressed in % of the row (label). Example 95.7% of all characters in the ZL transliteration have the same STA character as those in the FG transliteration (light green cell).