

Problème Pratique de Statistique - 20

Marqueurs binaires et chromosome Y

Les données sont publiées dans Semino *et al.* (2000). Un total de **1007** *Homo sapiens sapiens* d'Europe et des régions voisines, répartis en **25** groupes ont fourni un chromosome Y génotypé sur **19** marqueurs binaires (séquences de la région non-recombinante, marqueurs dits NRY de la région MSY pour *Male specific*)

Les données sont les fréquences (en 1 pour mille) des haplotypes rencontrés : pour chacun des marqueurs on a deux allèles et on dispose de la fréquence de l'un des deux (la fréquence de l'autre étant le complément à l'unité). La colonne `populations` donne le nom du groupe et est en général explicite. Sami désigne un peuple qui vit sur l'ensemble de l'espace arctique scandinave. Udmunt et Mari désigne des ethnies de deux Républiques de la fédération de Russie à l'est de Moscou.



`n` désigne le nombre d'échantillons du groupe, `Eu*` est le nom de code du marqueur défini en détail dans la légende et le fichier http://hpgl.stanford.edu/publications/Science_2000_22_Y_markers.doc. Un fond de carte est disponible ci-dessous. Les données sont accessibles par :

```
w = read.table("http://pbil.univ-lyon1.fr/R/pps/pps020.txt",h=T,r=1,as.is=T,sep="\t")
```

populations	n	Eu1	Eu3	Eu4	Eu6	Eu7	Eu8	Eu9	Eu10	Eu11	Eu12	Eu13	Eu14	Eu15	Eu16	Eu17	Eu18	Eu19	Eu20	Eu21
Andalusian	29	0	0	103	0	34	0	69	34	0	0	0	0	69	0	34	655	0	0	0
Spanish Basque	45	0	0	22	0	22	44	0	22	0	0	0	0	0	0	0	889	0	0	0
French Basque	22	0	0	0	0	0	91	45	0	0	0	0	0	0	0	0	864	0	0	0
Catalan	24	0	0	42	0	42	0	42	0	83	0	0	0	0	0	0	792	0	0	0
French	23	0	0	87	0	174	0	130	43	0	0	0	0	0	0	0	522	0	43	0
Dutch	27	0	0	37	0	222	0	0	0	0	0	0	0	0	0	0	704	37	0	0
German	16	0	0	62	0	375	0	0	0	0	0	0	0	0	0	0	500	62	0	0
Czech and Slovakian	45	0	0	22	0	156	0	89	0	44	0	0	22	22	22	0	356	267	0	0
Central-northern Italian	50	0	0	20	0	80	0	140	0	100	0	0	0	0	0	0	620	40	0	0
Calabrian	37	0	27	135	0	0	0	216	108	80	0	0	0	0	27	54	324	0	27	0
Sardinian	77	13	0	104	13	26	351	52	52	142	0	0	0	13	0	0	221	0	0	13
Croatian	58	0	0	69	0	448	0	52	0	17	0	0	0	0	17	0	103	293	0	0
Albanian	51	0	20	216	0	196	0	235	40	20	0	0	0	0	0	0	176	98	0	0
Greek	76	0	13	224	13	79	0	210	13	26	0	0	0	13	0	13	276	118	0	0
Macedonian	20	0	0	150	0	200	0	150	50	0	0	0	0	0	0	0	100	350	0	0
Polish	55	0	0	36	0	236	0	0	0	0	0	0	0	0	0	0	164	564	0	0
Hungarian	45	0	0	89	0	111	0	22	0	22	0	0	0	0	0	22	133	600	0	0
Ukrainian	50	0	0	40	0	180	0	60	0	40	20	0	60	20	0	0	20	540	20	0
Georgian	63	0	0	0	0	0	0	333	32	301	0	0	0	16	16	16	143	79	0	63
Turkish	30	0	33	133	0	33	0	400	33	66	0	0	33	33	33	33	66	66	33	0
Lebanese	31	0	0	258	32	32	0	290	161	32	0	0	0	0	0	32	64	97	0	0
Syrian	20	0	100	100	0	50	0	150	300	0	50	0	0	0	0	0	150	100	0	0
Saami	24	0	0	0	0	417	0	0	0	0	0	0	417	0	0	0	83	83	0	0
Udmurt	43	0	0	47	0	70	0	0	47	0	0	23	279	0	46	0	116	372	0	0
Mari	46	0	0	0	0	43	0	0	65	0	0	43	652	0	65	0	0	130	0	0

* The haplotypes are defined by the following markers and the respective derived alleles: Eu1, M13-C; Eu3, YAP+, 4064-A; Eu4, YAP+, 4064-A, M35-C; Eu6, RPS4-T; Eu7, M89-T, M170-C; Eu8, M89-T, M170-C, M26-A; Eu9, M89-T, M172-G; Eu10, M89-T; Eu11, M89-T, M201-T; Eu12, M89-T, M69-C; Eu13, M89-T, M9-G, TAT-C; Eu14, M89-T, M9-G, TAT-C, M178-T; Eu15, M89-T, M9-G, M70-C; Eu16, M89-T, M9-G; Eu17, M89-T, M9-G, M11-G; Eu18, M89-T, M9-G, M45-A, M173-C; Eu19, M89-T, M9-G, M45-A, M173-C, M17(delG); Eu20, M89-T, M9-G, M45-A; Eu21, M89-T, M9-G, M45-A, M124-T. Haplotypes Eu2, Eu5, and Eu22 were not observed.

Several samples were previously described (10, 11, 28). Samples not previously examined included 23 French, 16 Germans, 39 northern Italians, 45 Sardinians, 58 Croats, 20 Macedonians from northern Greece, 55 Poles,



La question est de savoir si ce type de marqueurs restitue une information spatiale concordante avec les nombreuses analyses sur l'histoire d'*Homo sapiens sapiens* en Europe. Lire l'article est indispensable. On pourra analyser les résultats de l'ACP faite par les auteurs. Quel tableau et quel codage permet de faire une AFC ? Comparer les deux approches. Une réflexion entre ACP normée et AFC sur tableau dédoublé peut suivre.

Semino, O., G. Passarino, P. J. Oefner, A. A. Lin, S. Arbizova, L. E. Beckman, G. De Benedictis, P. Francalacci, A. Kouvatsi, S. Limborska, M. Marcikiae, A. Mika, B. Mika, D. Primorac, A. Santachiara-Benerecetti, L. L. Cavalli-Sforza, and P. A. Underhill. 2000. The Genetic Legacy of Paleolithic *Homo sapiens sapiens* in Extant Europeans: A Y Chromosome Perspective. *Science* **290**:1155-1159.