

Ferret: a user-friendly tool to extract data from the 1000 Genomes Project

S. Limou, A. Taverner, N. Vince, P.-A. Gourraud, C.-A. Winkler

► **To cite this version:**

S. Limou, A. Taverner, N. Vince, P.-A. Gourraud, C.-A. Winkler. Ferret: a user-friendly tool to extract data from the 1000 Genomes Project. EFI 2017, May 2017, Mannheim, Germany. inserm-02161754

HAL Id: inserm-02161754

<https://www.hal.inserm.fr/inserm-02161754>

Submitted on 21 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Ferret: a user-friendly tool to extract data from the 1000 Genomes Project

S. Limou¹⁻⁴, A. Taverner^{4,5}, N. Vince^{1,2}, PA. Gourraud^{1,2}, and CA. Winkler⁴

¹Centre de Recherche en Transplantation et Immunologie (CRTI) UMR 1064, INSERM, Nantes, France

²Institut de Transplantation Urologie Néphrologie (ITUN), CHU Nantes, Nantes, France

³Ecole Centrale de Nantes, Nantes, France

⁴Basic Research Laboratory, NCI/NIH, Leidos Biomedical Research Inc., Frederick MD, USA

⁵Quantitative and Computational Biology program, Princeton University, Princeton NJ, USA

Contact:
ferret@nih.gov

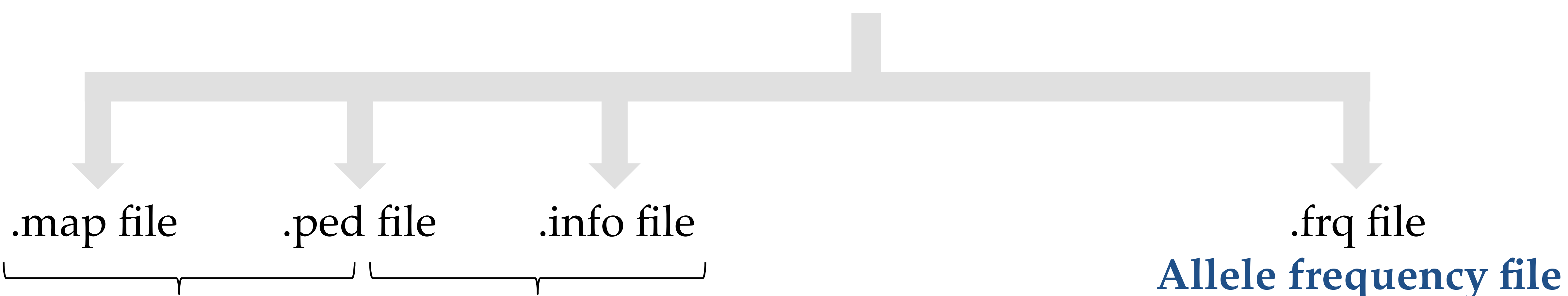
Additional settings:

- Genome version
- Output format
- Frequency filter

Input flexibility:

- Locus position
- Gene
- Variant ID

Retrieve 1000 Genomes Project data
(and optionally Exome Sequencing Project –ESP– data)



PLINK files

can be merged with user's dataset, etc.

HaploView files

data visualization, LD pattern, haplotypes, tagSNP design, etc.

CHR	VARIANT	POS	A1	A2	CHR	NB_1KG_1KG_A1_FREQ	ESP6500_EA_A1_FREQ	ESP6500_AA_A1_FREQ
6	rs1065711	31236853	G	A	1006	0.1213	.	.
6	rs1049853	31236900	G	A	1006	0.8867	0.8726	0.9209
6	indel_rs41548117_TC/T	31237100	TC	T	1006	0.9553	0.9594	0.9183
6	rs41542414	31237766	A	T	1006	1	0.9995	0.9897
6	rs9264623	31237950	T	C	1006	0.2773	.	.
6	indel_rs9281300_C/CA	31239170	C	CA	1006	0.4195	0.535	0.4125

Unique advantages of Ferret

- User-friendly interface
- Accepts input query as locus, gene(s), or variant(s)
- Handles SNPs and indels
- Outputs suitable for well-known pre-existing tools
- Computes allelic frequency for SNPs, indels and CNVs
- Retrieves allelic frequency from ESP

Perspectives:

- Minor GUI improvements
- Grant access to HLA alleles^s

Publicly available at <http://limousophie35.github.io/Ferret/>