FONDATION JEAN DAUSSET

Centre d'Étude du Polymorphisme Humain
Human Polymorphism Study Center

Président : André SYROTA

Directeur Scientifique : Jean François DELEUZE

Characteristics and access conditions to the HGDP-CEPH panel (2021-03-11)

Thank you for your interest in the HGDP-CEPH Human Genome Diversity Panel. In response to your request please find below the access conditions to the HGDP-CEPH panel

The main goal of the Panel is to further research in human population genetics. This is not a control panel for rare variants and samples cannot be used as controls for technological developments or validation.

A resource of 1063 lymphoblastoid cell lines (LCLs) from 1050 individuals in 52 world populations is presently stored at the Foundation Jean Dausset (CEPH), an academic, non-profit genome centre located in Paris. DNAs have been produced from these LCLs and organized into a panel at CEPH that is available for distribution to qualified, non-commercial, academic research laboratories (see table below).
The LCLs were collected from contributing research laboratories by the Human Genome Diversity Project (HGDP) and CEPH. They are listed in a table available on our website (http://www.cephb.fr/en/hgdp_panel.php). Some of these LCLs are also available in other collections. Information for each LCL is limited to sex of the individual and population and geographic origin. Identification is by code number only.

Although most of the LCLs originate from unrelated individuals, lines from first and second degree relative pairs have been identified in the panel. In addition, 13 of the 1063 LCLs making up the panel are duplicated, with one of the duplicate pairs reported to originate from different subpopulations; the duplicates are included in the panel as replicate controls. Furthermore, two individuals genotypically atypical for their reported populations of origin have been identified. Panel H952 contains no pairs of relatives closer than first cousins, no duplicate pairs and no atypical samples.

The LCLs of the HGDP-CEPH panel were produced by different laboratories in various countries over the past 20-30 years. In this project, DNA from these cell lines will be used by many laboratories in different countries. The HGDP-CEPH collaboration has determined that all of the blood samples used to produce these LCLs were collected with appropriate informed consent for the time and place of their collection. Recipients of DNA from these LCLs are, of course, responsible for ensuring that its use complies with legal standards that govern their laboratories. Both the legislation in your country in relation to the use of human biological samples and the General Data Protection Regulation (GDPR) must be respected.

Sixteen HGDP-CEPH LCLs show gender differences between those listed by the laboratories contributing the cell lines and those determined by molecular typing. Populations from which these 16 LCLs originated are indicated with asterisks in the table below, and the result of the genotyping is considered as the accurate gender.

Researchers who request the panel DNAs must send a summary of their project (one or two paragraphs) by email (HGDP_Manager@fjd-ceph.org). Genetic markers to be used should be described or preferably listed if practical; use official nomenclature and give their genome positions. For a resequencing study, indicate the genome region(s), the size of each region and how many individuals and corresponding populations to be sequenced.

Researchers must commit to type at least all DNAs of H952 with each marker used (at least 50 STRs, 50 common SNPs; for CNVs explain and describe the typing plan), and to communicate the results to CEPH, no later than 6 months after completion of typing the DNAs or at the time of publication (whichever comes first), for inclusion in a central database (http://www.cephb.fr/en/hgdp_panel.php#basedonnees), available on CEPH website. If these two conditions cannot be met please provide a scientific justification.

Collaborators must also agree by email (HGDP_Manager@fjd-ceph.org) to use the DNA samples for academic research only and not to transfer DNA samples to other laboratories without permission from CEPH.

If your project is approved, a quote will be sent to you. The HGDP-CEPH samples will be shipped after a purchase order from your purchase department is received by CEPH biobank manager (helene.blanche@fjd-ceph.org).
At the end of the project or on a deadline that will be fixed (usually 2 years after the shipment), researchers agree to destroy the HGDP-CEPH DNAs that they received for this specific project and to send a certificate of destruction to helene.blanche@fjd-ceph.org. The deadline can be postponed with CEPH approval.

In general DNAs, dissolved in TE (10:1), are aliquoted in tri-coded 0,5 mL FluidX tubes (2D code on base, 1D linear barcode and human readable number on side), at a concentration of ~60 ng/ul. The minimum quantity of DNA to be shipped will be ~5.0 micrograms per well. If you require more than 5.0 µg of each panel DNA, please contact us with the details of what you need.

We routinely send DNAs from the H952 panel (no "close relatives", no duplicate pairs) distributed in 952 tubes and placed into ten 2D barcoded racks.

The complete panel composed of 1163 samples (H1063) is also distributed. It is composed of the H952 panel plus 111 additional samples (13 DNAs from each individual of a duplicate pair excluded from H952, DNAs from the two atypical individuals (HGDP00980 and 00770), DNAs from 19 individuals each from a second degree relative pair excluded from H952 and DNAs from 77 individuals each from a first degree relative pair excluded from H952). These additional samples are distributed on two additional racks.

DNAs are experimental in nature, dedicated for research and may have hazardous properties. DNAs are delivered without any warranty. The recipient assumes all responsibilities for damages which may arise from the use, storage or disposal of received DNAs.

The HGDP-CEPH database gives you access to more than 100 billion allele frequencies calculated in 52 populations and 7 large population groups, representing more than 15 billion genomic variations. We recommend that you browse the database in order to learn how the HGDP-CEPH data might contribute to your research interests. You may find that markers of your interest have already been genotyped across the HGDP-CEPH panel.

The HGDP-CEPH panel was sequenced by the Wellcome Sanger Institute. The results were published in a Science paper on March 18, 2020: Global human genomes reveal rich genetic diversity shaped by complex evolutionary history. Sequencing data are available at https://www.ebi.ac.uk/ena/browser/view/PRJEB6463.

DNAs are made available worldwide upon request as a service to the scientific community. **DNAs must not be distributed to others without a written approval from CEPH Biobank manager.**

Researchers who wish to participate in the project as outlined above should send a summary of their project to HGDP_Manager@fjd-ceph.org and specifically indicate their agreement with the terms of collaboration (as above).

The DNA source will be cited in any publication reporting use of it in "Material and Methods" and "Acknowledgements" (as provided by CEPH Biobank, Paris, France [BIORESOURCES]) and publications will be sent to CEPH Biobank manager (Helene.Blanche@fjd-ceph.org).

As we do not have a specific budget to support managing the LCLs, DNA production, quality control and formatting, we must charge cost price for DNAs that you will receive. We can charge for DNA shipment or ship samples collect.

FONDATION JEAN DAUSSET

Centre d'Étude du Polymorphisme Humain
Human Polymorphism Study Center

Geographic origin of HGDP-CEPH samples

				H952 : LCL from unrelated (1rst and 2nd degree)			H1063 : All LCLs				
Region	Geographic Origin	Coordinates of Population samples	Population	Number of Males	Number of Females	Total number of LCL	Number of Males	Number of Females	Number of Male Duplicate	Number of Female Duplicate	Total number of LCL
Subsaharan Africa	Central African Republic	4N, 17E	Biaka Pygmy relatives	23		23	30	3	3		36
	Democratic Republic of Congo	1N, 29E	Mbuti Pygmy relatives	11	2	13	13	2			15
	Senegal	12N, 12W	Mandenka relatives	15	7	22	16	8			24
	Nigeria	6-10N, 2-8E	Yoruba relatives	12	10	22	13	12			25
	Namibia	21S, 20E	San relatives	6		6	7				7
	Kenya	3S, 37E	Bantu NE relatives	10	1	11	11	1			12
	S. Africa Bantu S.E	29S, 30E	Bantu S.E. Pedi	1		1	1				1
	S. Africa Bantu S.E	29S, 29E	Bantu S.E. Sotho	1		1	1				1
	S. Africa Bantu S.E	28S, 24E	Bantu S.E. Tswana	2		2	2				2
	S. Africa Bantu S.E	28S, 31E	Bantu S.E. Zulu	1		1	1				1
	S. Africa Bantu S.W.	22S, 19E	Bantu S.W. Herero	2		2	2				2
	S. Africa Bantu S.W.	19S, 18E	Bantu S.W. Ovambo	1		1	1				1
			Total Subsaharan Africa	**85**	**20**	**105**	**98**	**26**	**3**	**0**	**127**
North Africa	Algeria (Mzab)	32N, 3E	Mozabite relatives	20	9	29	20	10			30
			Total North Africa	**20**	**9**	**29**	**20**	**10**	**0**	**0**	**30**
Middle East	Israel (Negev)	31N, 35E	Bedouin relatives	27	19	46	28	20		1	49
	Israel (Carmel)	32N, 35E	Druze relatives	11	30	41	13	33		1	47
	Israel (Central)	32N, 35E	Palestinian relatives	16	30	46	17	34			51
			Total Middle East	**54**	**79**	**133**	**58**	**87**	**0**	**2**	**147**
Asia	Pakistan	30-31N, 66-67E	Brahui	25		25	25				25
	Pakistan	30-31N, 66-67E	Balochi relatives	24		24	25				25
	Pakistan	33-34N, 70E	Hazara relatives	23		23	25		1		26
	Pakistan	26N, 62-66E	Makrani	20	5	25	20	5			25
	Pakistan	24-27N, 68-70E	Sindhi relatives	20	4	24	21	4			25
	Pakistan	32-35N, 69-72E	Pathan	19	5	24	19	5			24
	Pakistan	35-37N, 71-72E	Kalash relatives	18	5	23	20	5			25
	Pakistan	36-37N, 73-75E	Burusho	20	5	25	20	5			25
	China	26-39N, 108-120E	Han	24	20	44	24	20		1	45
	China	29N, 109E	Tujia (minority)	9	1	10	9	1			10
	China	28N, 103E	Yizu (Yi) (minority)	9	1	10	9	1			10
	China	28N, 109E	Miaozu (Miao) (minority)	7	3	10	7	3			10
	China	48-53N, 122-131E	Oroqen (minority) relatives	6	3	9	7	3			10
	China	48-49N, 124E	Daur (minority)	7	3	10	7	3			10
	China	48-49N, 118-120E	Mongola (minority)	7	3	10	7	3			10
	China	47-48N, 132-135E	Hezhen (minority)	6	3	9	6	3	1		10
	China	43-44N, 81-82E	Xibo (minority)	8	1	9	8	1			9
	China	44N, 81E	Uygur (minority)	8	2	10	8	2			10
	China	21N, 100E	Dai (minority)	7	3	10	7	3			10
	China	22N, 100E	Lahu (minority) relatives	7	1	8	7	3			10
	China	27N, 119E	She (minority)	7	3	10	7	3			10
	China	26N, 100E	Naxi (minority) relatives	7	2	9	8	2			10
	China	36N, 101E	Tu (minority)	7	3	10	7	3			10
	Siberia	62-64N, 129-130E	Yakut	18	7	25	18	7			25
	Japan	38N, 138E	Japanese	21	8	29	22	8	1		31
	Cambodia	12N, 105E	Cambodian relatives	6	4	10	6	5			11
			Total Asia	**340**	**95**	435	**349**	**98**	**3**	**1**	**451**
Oceania	New Guinea	4S, 143E	Papuan	13	4	17	13	4			17
	Bougainville	6S, 155E	NAN Melanesian relatives	4	7	11	8	11		3	22
			Total Oceania	**17**	**11**	28	**21**	**15**	**0**	**3**	**39**
Europe	France	46N, 2E	French (various regions) relatives	12	16	28	12	17			29
	France	43N, 0	Basque	16	8	24	16	8			24
	Italy	40N, 9E	Sardinian	16	12	28	16	12			28
	Italy	46N, 10E	from Bergamo	8	5	13	8	5	1		14
	Italy	43N, 11E	Tuscan	6	2	8	6	2			8
	Orkney Islands	59N, 3W	Orcadian relatives	7	8	15	7	9			16
	Russia Caucasus	44N, 39E	Adygei	7	10	17	7	10			17
	Russia	61N, 39-41E	Russian	16	9	25	16	9			25
			Total Europe	**88**	**70**	158	**88**	**72**	**1**	**0**	**161**
America	Mexico	29N, 108W	Pima (relatives)	9	5	14	14	11			25
	Mexico	19N, 91W	Maya (relatives)	2	19	21	3	22			25
	Colombia	3N, 68W	Piapoco and Curripaco relatives	2	5	7	5	8			13
	Brazil	10S, 63W	Karitiana (relatives)	6	8	14	10	14			24
	Brazil	11S, 62W	Surui (relatives)	4	4	8	11	10			21
			Total America	**23**	**41**	64	**43**	**65**	**0**	**0**	**108**
			Total HGDP-CEPH LCLs	**627**	**325**	**952**	**677**	**373**	**7**	**6**	**1063**