

# An Introduction to HipSci cell lines

The Human Induced Pluripotent Stem Cells Initiative (HipSci) is generating a large, high-quality reference panel of human iPSC lines for the research community. These lines are created from tissue donations from both healthy volunteers and patients from particular rare disease communities.

This document and its associated video describe the cell lines generated by the HipSci project and how to find and access the lines for your research.

The HipSci project has successfully obtained samples from donors in 15 rare disease cohorts, to be used for the derivation of iPSCs. Our largest cohort by far is our normal healthy donor iPSCs. A list of our rare disease cohorts is shown below:

- Bardet-Biedl syndrome
- Alport syndrome
- Hereditary Cerebellar Ataxia
- Congenital Hyperinsulinia
- Kabuki syndrome
- Primary immune deficiency
- Bleeding and Platelet disorders
- Hypertrophic Cardiomyopathy
- Usher syndrome and Congenital eye defects
- Monogenic Diabetes
- Batten disease
- Hereditary Spastic Paraplegia
- Retinitis Pigmentosa
- Macular Dystrophy
- Childhood Neurology

The successfully derived iPSC lines from all 16 cohorts are or will be available to the scientific community, via two cell banks: ECACC and EBiSC.

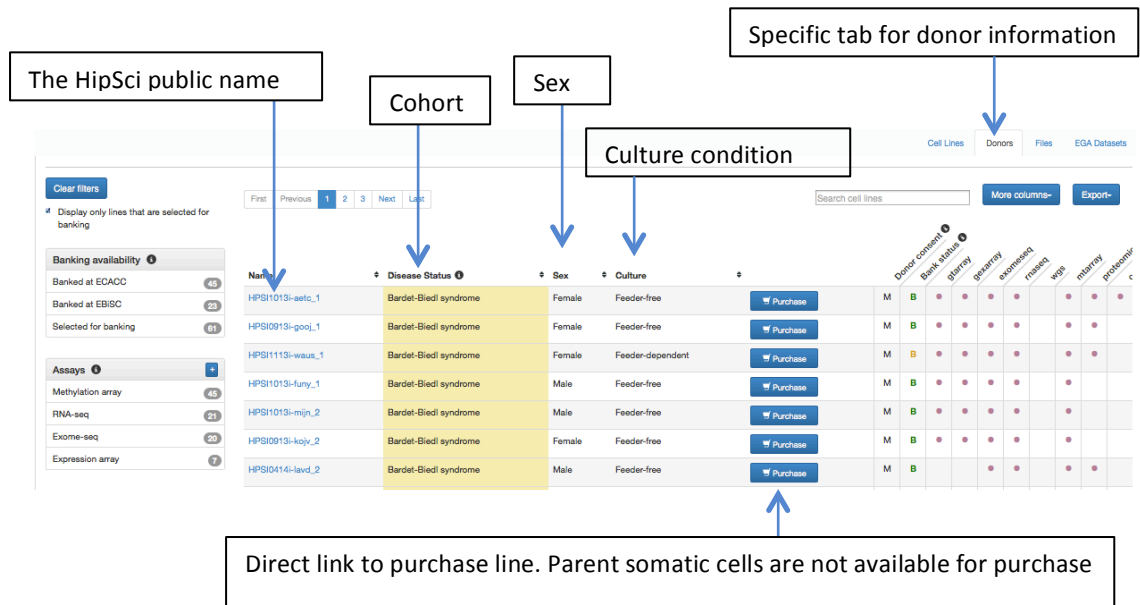
## Navigating the cell line browser

All the details of the cell lines HipSci derives are collated in our cell lines and data browser. You can launch the browser, via the push button at the top of the home, about us, cell line and data page called “Cell line and data browser”.

- 1) When you launch the cell line browser, a list of all our current ‘selected’ cell lines and their associated data is displayed. It is unlikely you will ever need to see all the lines (non-selected or in progress lines), but if you do, you can uncheck the box at the top left which is labelled “Display only lines that are selected for banking”.

☒ Display only lines that are selected for banking

- 2) The default cell line browser containing all the selected lines displays a wealth of information on the lines.



The HipSci public name

Cohort

Sex

Culture condition

Specific tab for donor information

Direct link to purchase line. Parent somatic cells are not available for purchase

Name	Disease Status	Sex	Culture	Purchase	Donor contact	Bank status	genotype	expression	transcriptome	single-cell	proteome
HPSI1013i-aetio_1	Bardet-Biedl syndrome	Female	Feeder-free	<a href="#">Purchase</a>	M	B					
HPSI0913i-goof_1	Bardet-Biedl syndrome	Female	Feeder-free	<a href="#">Purchase</a>	M	B					
HPSI1113i-waue_1	Bardet-Biedl syndrome	Female	Feeder-dependent	<a href="#">Purchase</a>	M	B					
HPSI1013i-funyn_1	Bardet-Biedl syndrome	Male	Feeder-free	<a href="#">Purchase</a>	M	B					
HPSI1013i-mijn_2	Bardet-Biedl syndrome	Male	Feeder-free	<a href="#">Purchase</a>	M	B					
HPSI0913i-kojv_2	Bardet-Biedl syndrome	Female	Feeder-free	<a href="#">Purchase</a>	M	B					
HPSI0414i-lavd_2	Bardet-Biedl syndrome	Male	Feeder-free	<a href="#">Purchase</a>	M	B					

- 3) To help understand the remaining information contained in the cell line browser, it is important grasp what each column depicts and what the list of abbreviations shown mean:

### Donor consent: A reference to the category the data and cell line falls under

Lines from our normal healthy donors fall under either managed or open access, based on the ethics conditions they were collected under. All HipSci iPSC lines from donors with rare diseases fall under managed access.

For a managed access line, the recipient agrees that any genetic or genomic data it generates from use of the Cell Lines will be held securely and only used in biomedical research, and will only be made available to third party researchers under a Data Access Agreement at least as stringent as the Data Access Agreement for data from the same Cell Lines at the EGA submitted by Sanger. Sanger encourages submission of genomic data to a managed access repository such as the EGA or dbGAP, with access being restricted to use for biomedical research and subject to a Data Access Agreement at least as stringent as that for data from the same Cell Lines at the EGA submitted by Sanger. For the avoidance of doubt, genetic or genomic data from the Cell Lines may not be made available on the open Internet.

For an open access line, the recipient agrees not to make any attempt to identify the original donors of the Cell Lines. Recipient agrees to remove any data mapping to the Y chromosome (including sequencing reads, genotypes etc.) before making genetic or genomic data available to third parties. Once this has been done, data may be made available for biomedical research in an open access fashion, permitting redistribution of derived data sets.

### Banked status

- **Selected for banking (B):** Cell lines that have passed HipSci's QC standards, and will continue in the pipeline towards being frozen down and banked at ECACC or

EBiSC.

- **Banked at ECACC (B):** Cell lines that are currently available from the [European Collection of Cell Cultures](#)
- **Banked at ECACC and EBiSC (B):** Cell lines that are available from the [European Collection of Cell Cultures](#) and [European Bank for induced pluripotent stem cells](#)
- **Pending selection (P):** Cell lines yet to undergo QC, or cell lines awaiting a decision on whether they are selected. Typically, HipSci strives to derive two iPS cell lines for each donor, both of which will be selected for banking at a cell bank, if they pass QC (refer to step 1 for instructions on how to see 'Pending' lines).
- **Not selected (N):** Cell lines that will never be available from a cell bank. Typically, this means another cell line from the same donor has been preferentially selected, or the line has failed to pass QC (refer to step 1 for instructions on how to see 'Not selected' lines).

#### Gtarray=Genotyping array

HipSci assays for genotypes in all iPS cell lines and in the somatic cells from which they were derived (e.g. fibroblasts). HipSci's QC procedure compares the called genotypes between the primary somatic and derived iPS cells of the same donor, to assess genomic integrity in the derived lines.

#### Gexarray=Expression array

Transcription profiling by array is used to measure the activity of genes in the HipSci cell lines. HipSci assays the transcriptome expression profile in all iPS cell lines and their somatic cell equivalent. HipSci's QC procedure uses transcript expression to assess if the derived cell lines have a profile typical of the pluripotent state.

#### Exomeseq=Exome sequencing

Exome sequencing is performed on all HipSci iPS cell lines selected for banking after passing QC. We have also generated exome sequencing data for ~250 healthy donor somatic cells. Sequencing and primary analysis are performed at the Wellcome Trust Sanger Institute.

#### Rnaseq=RNA sequencing

RNA-sequencing is performed on all HipSci iPS cell lines selected for banking after passing QC. Sequencing and primary analysis are performed at the Wellcome Trust Sanger Institute.

#### Wgs=Whole genome sequencing

A small selection (~150 donors) of iPS cell lines and their primary somatic cells have been deep whole genome sequenced at 30x coverage.

#### Mtarray=Methylation array

Methylation profiling by array is used to probe the epigenetic state of HipSci iPS cell lines. HipSci assays iPS cell lines that are selected for banking after passing QC. This assay and primary analysis is performed at the Wellcome Trust Sanger Institute.

#### Proteomics

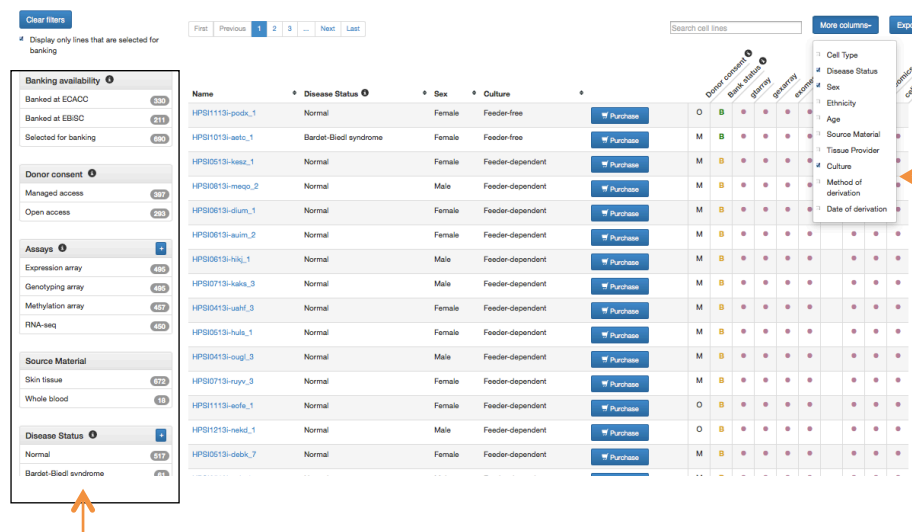
The proteomics work undertaken at The University of Dundee, includes mass spectrometry (MS)-based measurements of protein expression in many of the HipSci cell lines, from both normal and disease groups. This includes both label-free and TMT measurements of protein abundance. Work is on-going to closely integrate the quantitative measurements of protein expression levels with

parallel data generated within HipSci that document genome sequences and methylation patterns, mRNA expression and cell phenotypes, measured on the same cell lines.

### Cellbiol-fn=Cellular phenotyping

HipSci's cellular phenotyping assays are carried out at King's College London, at the Centre for Stem Cells & Regenerative Medicine. Imaging the HipSci cell lines and using high content analysis, we quantify and standardise phenotypic features that are informative of cell behaviour in an artificial microenvironment. Methods have been developed to facilitate the integration of dynamic and end-point imaging data, with genomics and proteomics data. This allows for the study of genetic contribution on cellular phenotype, and the association of single nucleotide changes to parameters such as cell adhesion, morphology and proliferation.

- It is also important to note there are other features in the browser that are not displayed by default, but can be displayed as shown below.



The screenshot shows the HipSci browser interface. On the left, there are filter panels for 'Banking availability', 'Donor consent', 'Assays', 'Source Material', and 'Disease Status'. The main table displays columns for Name, Disease Status, Sex, Culture, and a 'Purchase' button. To the right of the table, a 'More columns' dropdown menu is open, showing additional features like Cell Type, Disease Status, Sex, Ethnicity, Age, Source Material, Tissue Provider, Culture, Method of derivation, and Date of derivation. An orange arrow points to the 'More columns' button, and another orange arrow points to the 'Banking availability' filter panel.

Here you can filter by:

- \*Banking availability-ECACC, EBISC or hasn't been banked yet (only selected)
- \*Donor consent-managed or open
- \*Lines with a particular assay data type
- \*Source material-Material used to derive the iPSC line
- \*Disease status: Cohorts for which we have iPSC lines
- \*Age-Age band recipient falls into
- \*Tissue provider-Supplier of source material

The default browser, does not display all the details available per line. If you look under the 'more columns' button, you will see additional features can be displayed on the screen to help display a line of interest. These include, cell type (fibroblast, PBMC or iPSC line), age, ethnicity, source material, tissue provider, method of derivation and date of derivation. There are alternate options if you search for a dataset ID

## What information is available per line?

Each cell line displayed in our cell line browser can be explored further by clicking on the cell line name.

As way of example, below is a figure of what you will see if you click on the cell line name HPSI1113i-podx\_1. The sections have been divided to provide an explanation of each section.

1

# HPSI1113i-podx\_1

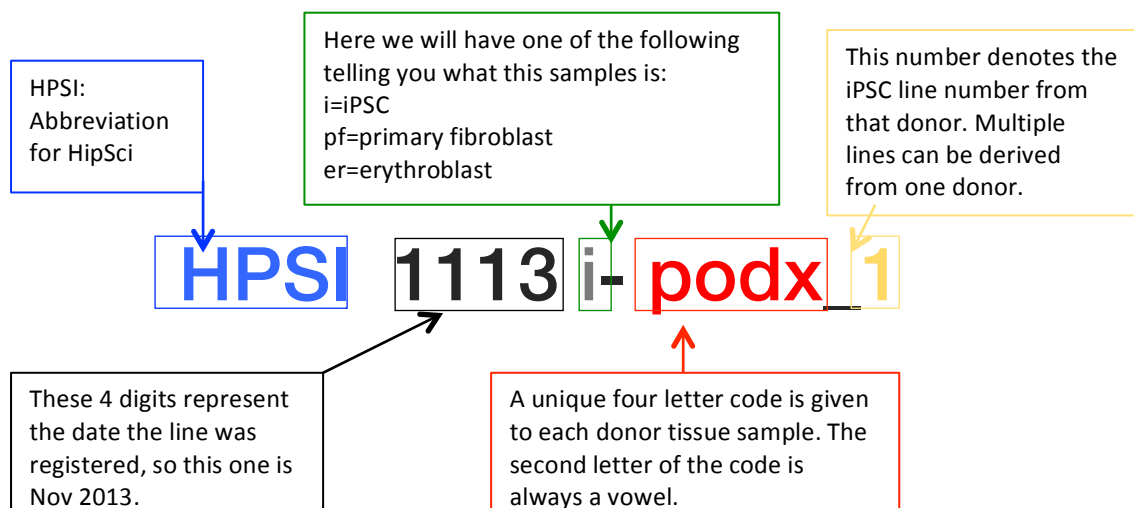
Purchase

## HipSci cell line


---

<b>Cell line type:</b>	iPSC	
<b>Banking status:</b>	<a href="#">Banked at ECACC</a>	2
	<a href="#">Banked at EBiSC as WTSII005-A</a>	3
<b>Donor:</b>	HPSI-podx	
<b>Disease status:</b>	Normal	4
<b>Source material:</b>	Skin tissue	
<b>Tissue provider:</b>	Cambridge BioResource	
<b>Method of derivation:</b>	CytoTune 1	
<b>Date of derivation:</b>	2014-03-13	
<b>Donor age:</b>	65-69	
<b>Donor ethnicity:</b>	White - White British	
<b>Donor sex:</b>	Female	
<b>Culture conditions:</b>	Feeder-free when banked	
<b>Biosample:</b>	<a href="#">SAMEA2533339</a>	5
<b>Certificate of analysis:</b>	<a href="#">CoA</a>	6
<b>SOP for culturing:</b>	<a href="#">Feeder-free SOP</a>	

- 1 Firstly, the cell line name itself will give you valuable information. You will see that each of our cell lines has a unique naming system. This standardised naming system was proposed by Luong *et al* in 2011, and was employed by HipSci.



- 2 The banking status links here will take you directly to the page where you can purchase the line itself. If you are using the lines for non-commercial research, then purchase the lines from [ECACC \(HipSci collection\)](#).
- 3 If you are intending to use them for commercial research, you can obtain the line from [EBiSC](#). Note: all lines available in the EBiSC catalogue are available to academic researchers via the [HipSci-ECACC collection](#) for a lesser price.
- 4 Pertinent cell line information
- 5 If you are looking to purchase a line, this CoA document will be very helpful to know specifics about the quality and status of the cell line. This CoA PDF displays the following information.



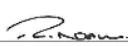

**Certificate of Analysis for HipSci iPSC**

Cell Line Name	HPS11131-podx_1	Culture and Passaging Methods.	Feeder Free*
Biosample ID	SAMBA2533339	Catalogue No.	77650036
Reprogramming Method	CytoTune® 1	Lot.	9.10.14
Disease Association	Normal	Donor Cell Material	Skin tissue
Gender	Female	Passage No.	p31
Associated Data and Publications	<a href="http://www.hipsci.org/lines/8/lines">http://www.hipsci.org/lines/8/lines</a> <a href="http://www.ebi.ac.uk/biosamples/browse_samples.html?keywords=hipsci">http://www.ebi.ac.uk/biosamples/browse_samples.html?keywords=hipsci</a>		


The following standard testing criteria have been determined within CGAP, prior to release of this product:

Test	Assay	Result
Confirmed Sterility	PCR for Mycoplasma	Pass
Cell Line Identity	Fluidigm	Pass
Viability post-thaw	Growth to confluence post-thaw	Pass
Morphology	Continuous visual assessment of iPSC colony morphology.	Pass
Stem Cell Marker Expression	Pluri test	<a href="http://www.hipsci.org/lines/8/lines/HPS11131-podx_1">http://www.hipsci.org/lines/8/lines/HPS11131-podx_1</a>
Clearance of Reprogramming Factors	rtPCR analysis	Pass

\*These Cell lines were cultured in media containing Pen/Strep.

Acceptable for release: Signed  Date 26/02/16

Project Lead

Agreed by: Signed  Date 26/2/16

Head of Operations

→ Culture type

→ No mycoplasma

→ Identity matches parent donor  
→ Resuscitation test status

→ Pluripotency status

→ Sendai clearance status

- 6 HipSci recommended protocol for culturing this particular iPSC line

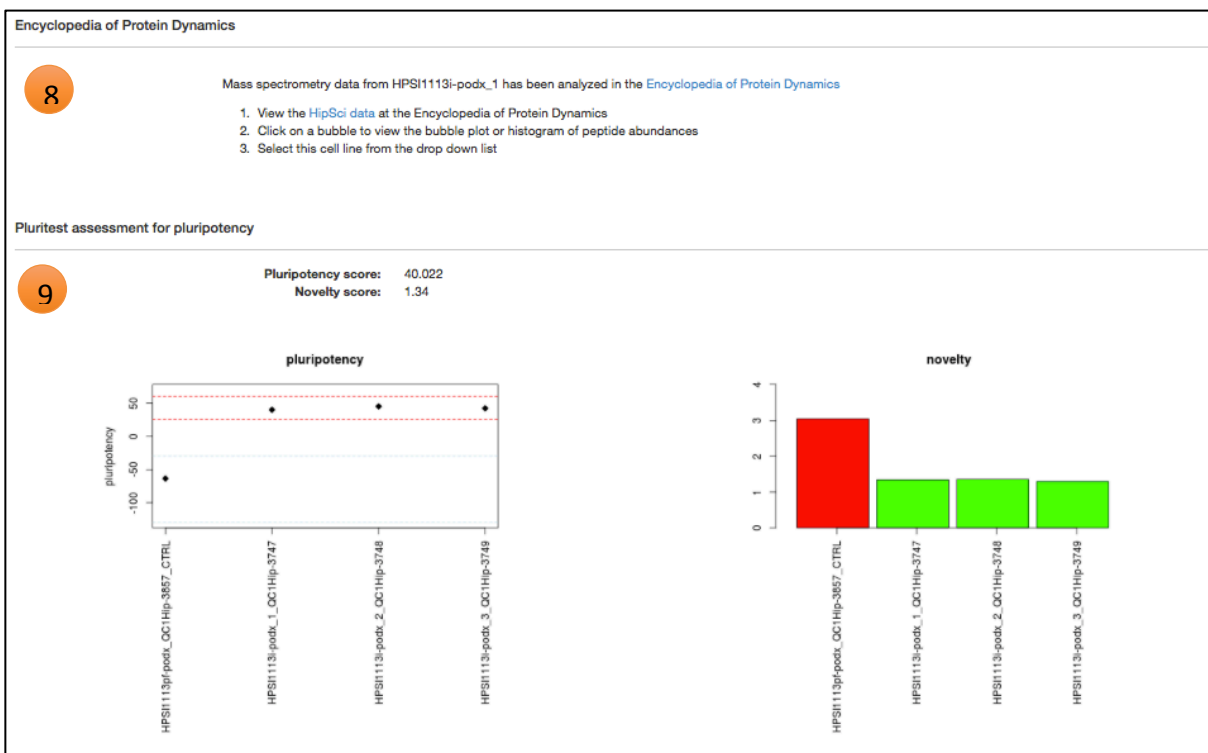
Genomic assay files

7

Export file list- More columns-

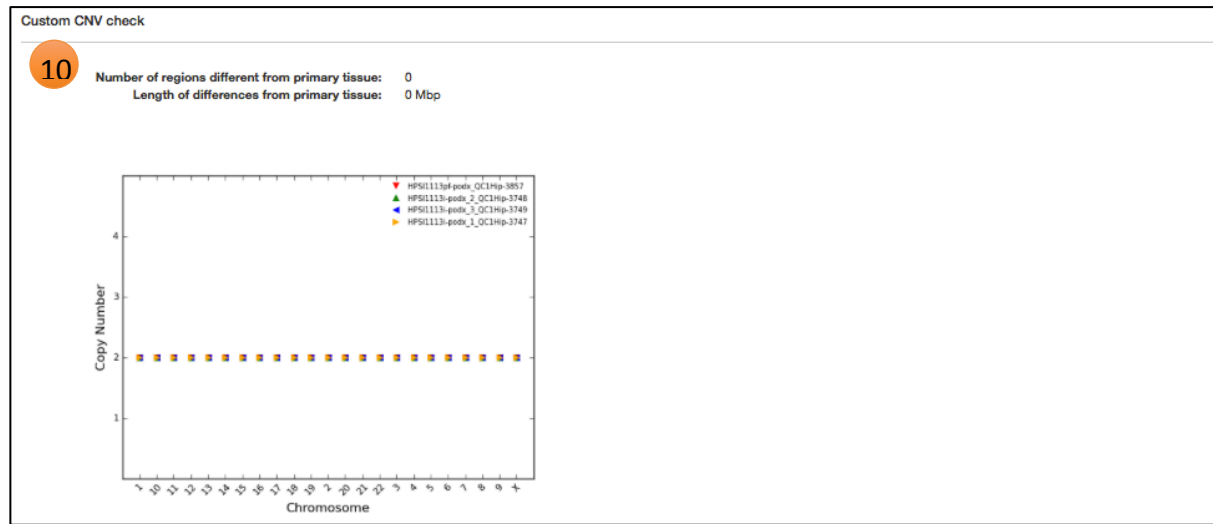
Assay	Description	File download	Archive	Culture	Passage
Exome-seq	Raw sequencing reads		ENA	Feeder-free	31
Exome-seq	mpileup variant calls		ENA	Feeder-free	31
Exome-seq	Imputed and phased genotypes		ENA	Feeder-free	31
Exome-seq	Raw sequencing reads		ENA	Feeder-free	31
Expression array	GenomeStudio text file		ArrayExpress	Feeder-dependent	12
Genotyping array	Genotyping array calls		ENA	Feeder-dependent	12
Genotyping array	Imputed and phased genotypes		ENA	Feeder-dependent	12
Genotyping array	HLA typing multiple cell lines		HipSci FTP	Feeder-dependent	
Methylation array	Text file with probe intensities		ArrayExpress	Feeder-free	31
RNA-seq	Raw sequencing reads		ENA	Feeder-free	31
RNA-seq	Raw sequencing reads		ENA	Feeder-free	31
RNA-seq	Abundances of transcripts		ENA	Feeder-free	31
RNA-seq	Splice-aware STAR alignment		ENA	Feeder-free	31
Whole genome sequencing	Raw sequencing reads		ENA	Feeder-free	31

7 Here you will see all the assay files currently available for download, either from ENA or EGA. Note: Some lines were produced as feeder-dependent lines and transferred to feeder-free later on in the pipeline. For this reason the data collected in these lines can be a mixture, some data on the feeder-dependent version and some the feeder-free version of the same line. In this example you will see that the expression array and genotyping array were done on the feeder-dependent version of the line, but all other assays were done on the feeder-free version.



8 The proteomics data that has been generated on this line can be easily visualised using the [Encyclopedia of Protein Dynamics](#).

9 These are the results from the expression array used to pass and select this line.



10 These are the results from the expression array used to pass and select this line.

## How do I know what is a good line to pick?

If you have a criterion in mind for the cell line you wish to obtain, you can manipulate the cell line browser to fish out the line you need. However, for new users of our iPSC resource, we have composed two panels (feeder-dependent and feeder-free) of lines from the HipSci iPSC collection that we would recommend. These lines have been selected based on their high pluritest scores, low CNV's, comprehensively characterised assay data, and their differentiation potential.

To display these panels, there are two options. You can click the "Recommended cell line panel" black box on the main cell page under cohorts, or launch the cell line and data browser from any of our pages. Once the browser is up, you will see the push button "Recommended cell line panel" at the very top of the cell line browser.

Recommended panel of HipSci lines

Feeder-free panel	
HPSI0214i-wibj_2	<a href="#">Purchase</a>
HPSI0214i-kucg_2	<a href="#">Purchase</a>
HPSI0314i-hoik_1	<a href="#">Purchase</a>
HPSI0314i-sojd_3	<a href="#">Purchase</a>

Feeder-dependent panel	
HPSI0613i-zisa_3	<a href="#">Purchase</a>
HPSI0813i-voas_2	<a href="#">Purchase</a>
HPSI0513i-cuau_2	<a href="#">Purchase</a>
HPSI0713i-kaks_3	<a href="#">Purchase</a>



If you are looking for the associated assay data for these lines, simply click on the line name and it will take you to the data files you can download.

## Where to get help

If you have a query about the availability of a line or need help with ordering, you can email [culturecollections.technical@phe.gov.uk](mailto:culturecollections.technical@phe.gov.uk)

If you are struggling to find the information you need regarding our cell lines and data, or have another general enquiry, you can contact us at [hipsci@ebi.ac.uk](mailto:hipsci@ebi.ac.uk)

## Acknowledgements

Thanks to Ian Streeter, Laura Clarke and Reena Halai for producing this document.