

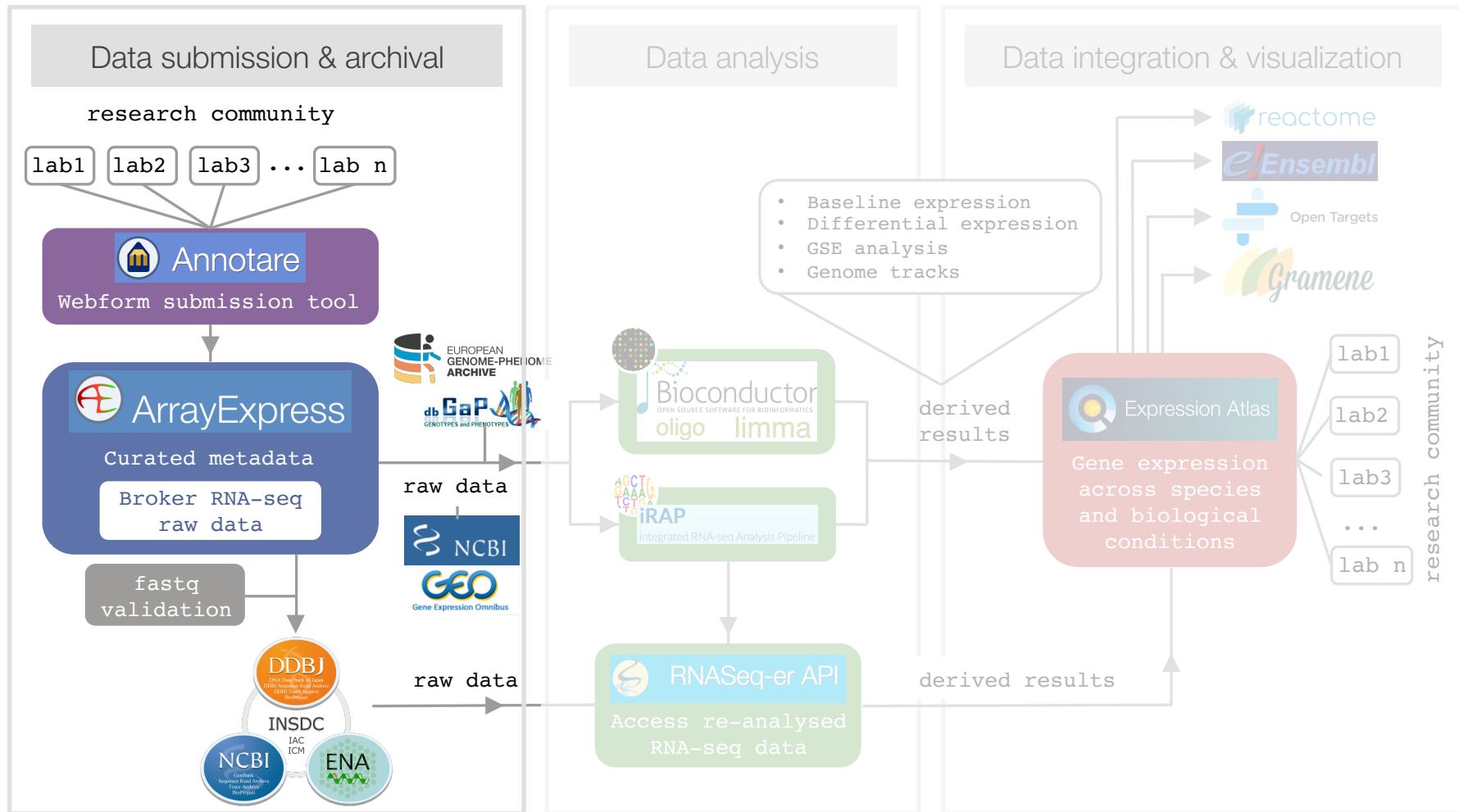
EMBL-EBI workshop: Data and tools for transcriptomics and protein biology

Discover functional genomics data
with ArrayExpress

Laura Huerta, PhD
Senior Scientific Curator
lauhuema@ebi.ac.uk
Malta, 31 July 2018



Functional genomics resources at EMBL-EBI





ArrayExpress: functional genomics data archive

 ArrayExpress

Home Browse Submit Help About ArrayExpress Contact Us Login

Search Examples: E-MEXP-31, cancer, p53, Geuvadis [advanced search](#)

ArrayExpress – functional genomics data

ArrayExpress Archive of Functional Genomics Data stores data from high-throughput functional genomics experiments, and provides these data for reuse to the research community.

 [Browse ArrayExpress](#)

Latest News

18 May 2018 - **New single-cell submission and metadata guide**
We are happy to announce the newest addition to our help pages: [The single-cell submission guide!](#)
This is a simple step-by-step guide which leads you through submitting single-cell sequencing submissions and lists the minimum required metadata to include. The guide starts by explaining the expected sample attributes dependent on the experiment design to help you correctly annotate your samples. The next section details the technical information that is required for single cell-specific procedures including pre/post analysis quality control or how single cells were generated. The library and sequencing information section helps you to understand the required information for the library construction protocols, how to include barcode information, and spike-in information where applicable. Lastly, how to prepare the raw data files for your submission.
We hope you find this useful to submit single-cell sequencing data!

Links

Information about how to search ArrayExpress, understand search results, how to submit data and FAQ can be found in our [Help section](#).

Find out more about the [Functional Genomics group](#).

Tools and Access

[Annotare](#): web-based submission tool for ArrayExpress.

[ArrayExpress Bioconductor package](#): an R package to access ArrayExpress and build data structures.

[Programmatic access](#): query and download data using web services or JSON.

Related Projects

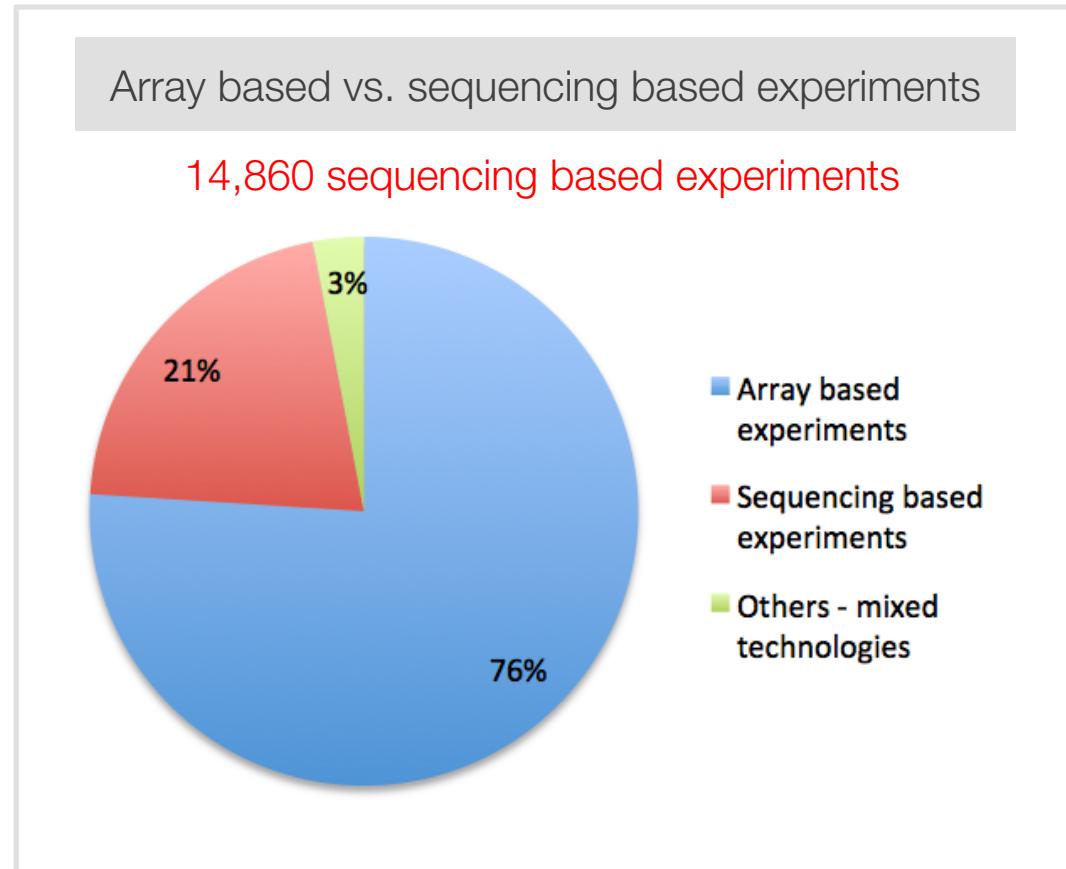
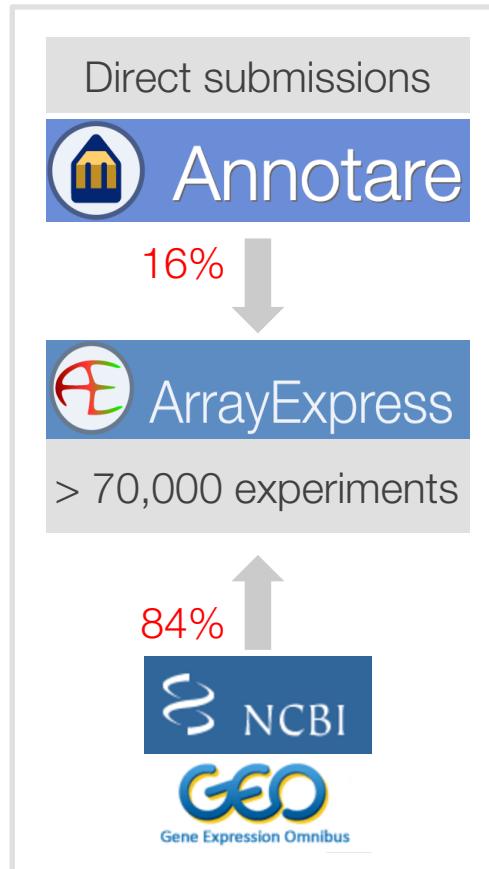
Discover up and down regulated genes in numerous experimental conditions in the [Expression Atlas](#).

Explore the [Experimental Factor Ontology](#) used to support queries and annotation of ArrayExpress data.

www.ebi.ac.uk/arrayexpress



ArrayExpress: functional genomics data archive

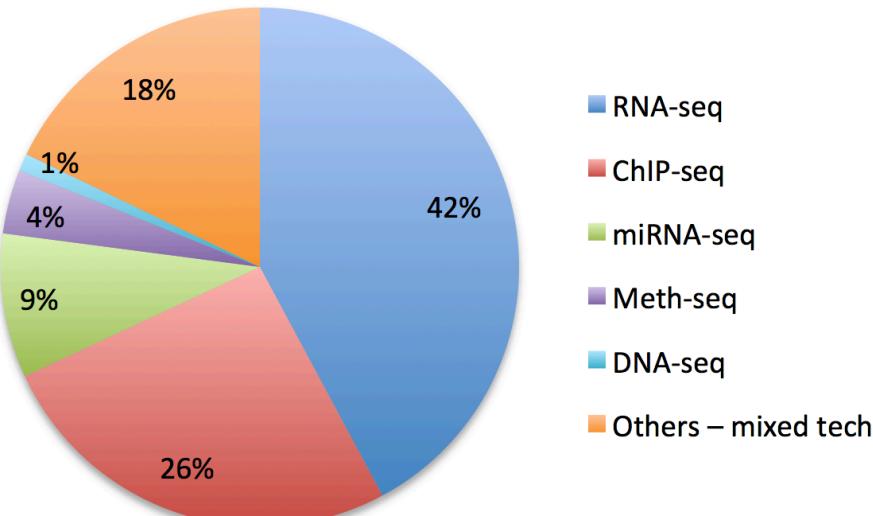




ArrayExpress: functional genomics data archive

Sequencing based experiments

6,266 RNA-seq of coding RNA experiments





ArrayExpress: functional genomics data archive

[Home](#)[Browse](#)[Submit](#)[Help](#)[About ArrayExpress](#)

Experiment types in ArrayExpress

Experiment Type	Definition (EFO term accession)	Example
4C	EFO_0007690	E-MTAB-2180
antigen profiling	EFO_0000747	E-MTAB-3606
ATAC-seq	EFO_0007045	E-MTAB-3972
Bisulfite-seq	EFO_0003753	E-MTAB-1042
Capture-C	EFO_0007691	E-MTAB-4845
ChIP-chip by array	EFO_0002760	E-MTAB-2804
ChIP-chip by SNP array	EFO_0002764	E-GEOID-22306
ChIP-chip by tiling array	EFO_0002762	E-MTAB-1402
ChIP-seq	EFO_0002692	E-MTAB-3631
CLIP-seq	EFO_0003143	E-MTAB-1371
comparative genomic hybridization by array	EFO_0000749	E-MTAB-2293
DNA-seq	EFO_0002693	E-MTAB-3109
FAIRE-seq	EFO_0004428	E-MTAB-3199

www.ebi.ac.uk/arrayexpress/help/experiment_types.html



Data mining with ArrayExpress

The screenshot shows the ArrayExpress homepage. At the top, there's a search bar with examples like "E-MEXP-31, cancer, p53, Geuvadis" and a link to "advanced search". Below the search bar are links for "Contact Us" and "Login". The main content area features a large yellow starburst graphic. On the left, there are links for "Home", "Browse" (which is highlighted in yellow), "Submit", "Help", and "About ArrayExpress". The central part of the page displays the title "ArrayExpress – functional genomics data" and a brief description: "ArrayExpress Archive of Functional Genomics Data stores data from high-throughput functional genomics experiments, and provides these data for reuse to the research community." Below this is a blue button labeled "Browse ArrayExpress". A "Latest News" section follows, featuring a post about a new single-cell submission guide. The news post includes a date ("18 May 2018"), a summary, and a link to the guide. The bottom of the news section notes that users hope it will be useful for single-cell sequencing submissions.

🔗 Links

Information about how to search ArrayExpress, understand search results, how to submit data and FAQ can be found in our [Help section](#).

Find out more about the [Functional Genomics group](#).

🔧 Tools and Access

[Annotare](#): web-based submission tool for ArrayExpress.

[ArrayExpress Bioconductor package](#): an R package to access ArrayExpress and build data structures.

[Programmatic access](#): query and download data using web services or JSON.

🔗 Related Projects

Discover up and down regulated genes in numerous experimental conditions in the [Expression Atlas](#).

Explore the [Experimental Factor Ontology](#) used to support queries and annotation of ArrayExpress data.

www.ebi.ac.uk/arrayexpress



Data mining with ArrayExpress

ArrayExpress

Search Examples: E-MEXP-31, cancer, p53, Geuvadis advanced search

Home Browse Submit Help About ArrayExpress Contact Us Login

Filter search results

Page 1 2 3 4 5 6 .. 2847 Showing 1 - 25 of 71175 experiments

Sortable headings

Accession	Title	Type	Organism	Assays	Released	Processed	Raw	Views	Atlas
E-MTAB-6201	Small RNA-seq analysis of single porcine blastocysts revealed that maternal estradiol-17beta exposure does not affect miRNA isoform (isomiR) expression	microRNA profiling by high throughput sequencing	Sus scrofa	36	Yesterday			7	-
E-MTAB-5876	24h time-course microarray study of SW480 and SW620 cells	transcription profiling by array	Homo sapiens	18	Yesterday	-		5	-
E-MTAB-6236	Various evolutionary trajectories lead to loss of activity of tobramycin-potentiating compounds in Burkholderia cenocepacia biofilms	DNA-seq	Burkholderia cenocepacia	12	22/07/2018	-		12	-
E-MTAB-6452	Transcriptome analysis of HeLa cells upon deprivation of methionine and cysteine (Met/Cys)	RNA-seq of coding RNA	Homo sapiens	8	21/07/2018	-		16	-
E-MTAB-6894	Effect of the IRE1a-XBP1 pathway inhibition during T helper 2 differentiation	RNA-seq of coding RNA	Mus musculus	8	20/07/2018	-		-	-
E-MTAB-5934	Transcription profiling by array of B lymphoblasts derived from either control or patients with inactivation mutations in the gene for the transcription factor NF-κB1	transcription	Homo sapiens	10	19/07/2018	-		-	-

Accession number

Experiment type

Number of hybridisations, sequencing libraries



Data mining with ArrayExpress

Browse ArrayExpress

Hands-on activity

Find the biggest experiments (in terms of number of assays). Now display the top 100 most viewed experiments on one page



www.ebi.ac.uk/arrayexpress/experiments/browse.html



ArrayExpress: Experimental variable

Experimental variable



It is the main factor that you are investigating

Transcription profiling of blood from smokers, non-smokers and former smokers to identify gene expression signature for cigarette smoke exposure response

organism > Homo sapiens

age > 22, 57, 43, 39 year, etc.

sex > male OR female

organism part > blood

disease > normal

ethnic group > Caucasian OR African American

clinical history > smoker, former smoker OR non-smoker



ArrayExpress: Experimental variable

Experimental variable



It is the main factor that you are investigating

Transcription profiling of blood from smokers (with or without COPD), non-smokers and former smokers to identify gene expression signature for cigarette smoke exposure response

organism > Homo sapiens

age > 22, 57, 43, 39 year, etc.

sex > male OR female

organism part > blood

ethnic group > Caucasian OR African American

disease > normal OR COPD

clinical history > smoker, former smoker OR non-smoker



ArrayExpress: Experimental variable

Experimental variable

Hands-on activity

Go to experiment E-GEO-49515:

- 1) What is the experimental variable?
- 2) What groups of samples are compared?



www.ebi.ac.uk/arrayexpress



ArrayExpress: Experimental variable

	A	B	C
1	Source Name	Array Data File	FactorValue [disease]
2	GSM1200316 1	GSM1200316_PAN-PBMC-S03.CEL	pancreatic cancer
3	GSM1200315 1	GSM1200315_PAN-PBMC-S02.CEL	pancreatic cancer
4	GSM1200314 1	GSM1200314_PAN-PBMC-S01.CEL	pancreatic cancer
5	GSM1200313 1	GSM1200313_NOR-PBMC-S10.CEL	normal
6	GSM1200312 1	GSM1200312_NOR-PBMC-S09.CEL	normal
7	GSM1200311 1	GSM1200311_NOR-PBMC-S08.CEL	normal
8	GSM1200310 1	GSM1200310_NOR-PBMC-S07.CEL	normal
9	GSM1200309 1	GSM1200309_NOR-PBMC-S06.CEL	normal
10	GSM1200308 1	GSM1200308_NOR-PBMC-S05.CEL	normal
11	GSM1200307 1	GSM1200307_NOR-PBMC-S04.CEL	normal
12	GSM1200306 1	GSM1200306_NOR-PBMC-S03.CEL	normal
13	GSM1200305 1	GSM1200305_NOR-PBMC-S02.CEL	normal
14	GSM1200304 1	GSM1200304_NOR-PBMC-S01.CEL	normal
15	GSM1200303 1	GSM1200303_HCC-PBMC-S10.CEL	hepatocellular carcinoma
16	GSM1200302 1	GSM1200302_HCC-PBMC-S09.CEL	hepatocellular carcinoma
17	GSM1200301 1	GSM1200301_HCC-PBMC-S08.CEL	hepatocellular carcinoma
18	GSM1200300 1	GSM1200300_HCC-PBMC-S07.CEL	hepatocellular carcinoma
19	GSM1200299 1	GSM1200299_HCC-PBMC-S06.CEL	hepatocellular carcinoma
20	GSM1200298 1	GSM1200298_HCC-PBMC-S05.CEL	hepatocellular carcinoma
21	GSM1200297 1	GSM1200297_HCC-PBMC-S04.CEL	hepatocellular carcinoma
22	GSM1200296 1	GSM1200296_HCC-PBMC-S03.CEL	hepatocellular carcinoma
23	GSM1200295 1	GSM1200295_HCC-PBMC-S02.CEL	hepatocellular carcinoma
24	GSM1200294 1	GSM1200294_HCC-PBMC-S01.CEL	hepatocellular carcinoma
25	GSM1200293 1	GSM1200293_GAS-PBMC-S03.CEL	gastric cancer
26	GSM1200292 1	GSM1200292_GAS-PBMC-S02.CEL	gastric cancer
27	GSM1200291 1	GSM1200291_GAS-PBMC-S01.CEL	gastric cancer



Data mining with ArrayExpress

Data mining: Discover datasets to analyse yourself

Find RNA-seq experiments related human autoimmune disease

main search box

The screenshot shows the ArrayExpress homepage with a blue header. On the left is the logo and the text "ArrayExpress". Below the header are navigation links: Home, Browse (which is highlighted in white), Submit, Help, About ArrayExpress, Contact Us, and Login. A yellow button labeled "Filter search results" is positioned below the navigation bar. To the right is a search bar with the placeholder "Search" and a magnifying glass icon. Below the search bar are examples of search terms: "E-MEXP-31, cancer, p53, Geuvadis" and a link to "advanced search". A red box highlights the search bar area, and another red box highlights the "Filter search results" button.

Search advanced search

Examples: E-MEXP-31, cancer, p53, Geuvadis

Filter search results

www.ebi.ac.uk/arrayexpress/browse.html



Data mining with ArrayExpress

Data mining: Discover datasets to analyse yourself

Find RNA-seq experiments related human autoimmune disease

Filter search results

By organism:

Homo sapiens

By experiment type:

RNA assay

Sequencing assay

By array:

All arrays

ArrayExpress data only

[Reset filters](#)

[Filter](#)

"autoimmune dis

autoimmune disease

- Autoimmune Hepatitis
- Behcet's syndrome
- CNS demyelinating **autoimmune disease**
- Eosinophilia-Myalgia Syndrome
- Guillain-Barre syndrome
- Myasthenia gravis
- Sjogren syndrome
- Susac Syndrome
- Vitiligo
- Wegener's granulomatosis
- anti-neutrophil antibody associated vasculitis
- antiphospholipid antibodies
- arthritis**
- autoimm

Efficient search via
ontology-driven query
expansion



Ontology-driven query expansion

Ontology Lookup Service

Home | **Ontologies** | Documentation | About

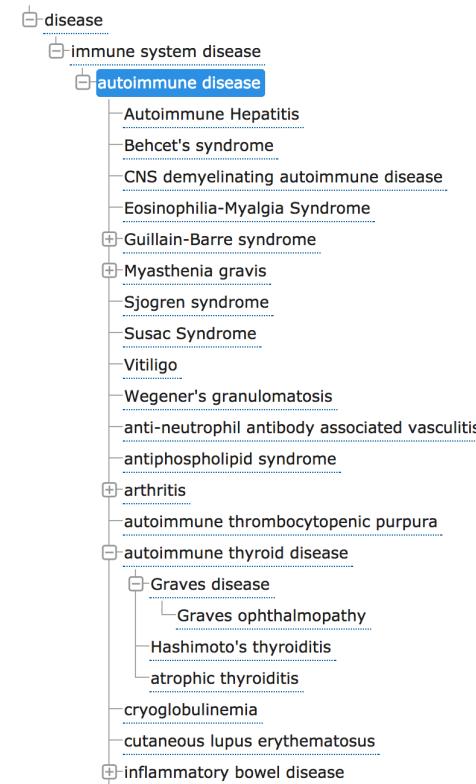
OLS > Experimental Factor Ontology > EFO:0005140

autoimmune disease

http://www.ebi.ac.uk/efo/EFO_0005140

Much richer queries by
using the hierarchy within
the ontology

Return subclasses when a
higher level superclass is
used in the query



www.ebi.ac.uk/ols/ontologies/efo



Data mining with ArrayExpress

Data mining: Discover datasets to analyse yourself

Find RNA-seq experiments related human autoimmune disease

Search results for "autoimmune disease"

Filtered by organism **Homo sapiens**, experiment type "sequencing assay", experiment type "rna assay"

Page 1 2

Showing 1 - 25 of 32 experiments

Page size 25 50 100 250 500

Accession	Title	Type	Organism	Assays	Released	Processed	Raw	Views	Atlas
E-GEOD-60424	Next generation sequencing of human immune cell subsets across diseases	RNA-seq of coding RNA	Homo sapiens	134	06/01/2015	-		3742	
E-MTAB-1568	RNA-seq of coding RNA in human plasma cells isolated from individuals with colorectal cancer, ulcerative colitis and normal individuals	RNA-seq of coding RNA	Homo sapiens	9	29/03/2013	-		1157	-
E-GEOD-46579	A blood based 12-miRNA signature of Alzheimer patients	RNA-seq of non coding RNA	Homo sapiens	70	17/06/2013	-		939	-
E-GEOD-57945	RNA-seq of 359 treatment-naive pediatric patients with Crohn's disease, patients with ulcerative colitis and control individuals	RNA-seq of coding RNA	Homo sapiens	359	24/07/2014			786	-
E-GEOD-83139	Single cell RNA-seq of human pancreatic endocrine cells from Juvenile, adult control and type 2 diabetic donors	RNA-seq of coding RNA	Homo sapiens	635	05/07/2016	-		436	-



Data mining with ArrayExpress

Samples
and files

Experiment
description

Experiment
type

MINSEQE
compliance score

Links to
data

E-GEOD-60424 - Next generation sequencing of human immune cell subsets across diseases

Status *Released on 6 January 2015, last updated on 25 February 2016*

Organism **Homo sapiens**

Samples (134) [Click for detailed sample information and links to data](#)
↳ found inside: multiple sclerosis, Type 1 Diabetes

Protocols (2) [Click for detailed protocol information](#)

Description This study compared whole transcriptome signatures of 6 immune cell subsets and whole blood from patients with an array of immune-associated diseases. Fresh blood samples were collected from healthy subjects and subjects diagnosed **type 1 diabetes**, amyotrophic lateral sclerosis, and sepsis, as well as **multiple sclerosis** patients before and 24 hours after the first treatment with IFN-beta. At the time of blood draw, an aliquot of whole blood was collected into a Tempus tube (Invitrogen), while the remainder of the primary fresh blood sample was processed to highly pure populations of neutrophils, monocytes, B cells, CD4 T cells, CD8 T cells, and natural killer cells. RNA was extracted from each of these cell subsets, as well as the whole blood samples, and processed into RNA sequencing (RNAseq) libraries (Illumina TruSeq). Sequencing libraries were analyzed on an Illumina HiScan, with a target read depth of ~20M reads. Reads were demultiplexed, mapped to human gene models (ENSEMBL), and tabulated using HTSeq. Read count data were normalized by the TMM procedure (edgeR package). We performed whole genome RNAseq profiling of immune cell subsets and whole blood from subjects with an array of immune-associated diseases.

Experiment type **RNA-seq of coding RNA**

Contacts Scott Presnell <SPresnell@benaroyaresearch.org>, Carla J Greenbaum, Cate Speake, Damien Chaussabel, Elizabeth Whalen, Jane H Buckner, Kimm K O'Brien, Michael J Mason, Peter S Linsley, Quynh-Anh Nguyen, Scott R Presnell, Uma Malhotra, Vivian H Gersuk

Citation [Copy number loss of the interferon gene cluster in melanomas is linked to reduced T cell infiltrate and poor patient prognosis.](#) Linsley PS, Speake C, Whalen E, Chaussabel D. , [Europe PMC 25314013](#)

MINSEQE * * * - *

Exp. design Protocols Variables Processed Seq. reads

Files Investigation description [↓ E-GEOD-60424.idf.txt](#)
Sample and data relationship [↓ E-GEOD-60424.sdrf.txt](#)
Additional data (1) [↓ E-GEOD-60424.additional.1.zip](#)

Links [Click to browse all available files](#)

Expression Atlas - E-GEOD-60424

ENA - SRP045500, GEO - GSE60424

Send E-GEOD-60424 data to



Data mining with ArrayExpress

E-GEOD-60424 - Next generation sequencing of human immune cell subsets across diseases

[Display full sample-data table](#)

[Export table in Tab-delimited format](#)

Page 1 2 3 4 5 6 .. 11

Showing 1 - 25 of 268 rows

Page size 25 50 100 250 500

Source Name	name	Sample Attributes							cell type	disease	ENA	FASTQ
		age	cell type	disease	individual	organism	ethnic group	sex				
GSM1479433 1	32 (year)	whole blood	normal	44	Homo sapiens	hispanic	female		whole blood	norm	EN	FT
GSM1479433 1	32 (year)	whole blood	normal	44	Homo sapiens	hispanic	female		whole blood	norm	EN	FT
GSM1479434 1	52 (year)	whole blood	multiple sclerosis	31	Homo sapiens	white	female	smoker	whole blood	multi	EN	FT
GSM1479434 1	52 (year)	whole blood	multiple sclerosis	31	Homo sapiens	white	female	smoker	whole blood	multi	EN	FT
GSM1479435 1	52 (year)	whole blood	multiple sclerosis	33	Homo sapiens	white	female	smoker	whole blood	multi	EN	FT
GSM1479435 1	52 (year)	whole blood	multiple sclerosis	33	Homo sapiens	white	female	smoker	whole blood	multi	EN	FT
GSM1479436 1	24 (year)	whole blood	Type 1 Diabetes	34	Homo sapiens	white	female		whole blood	Type	EN	FT
GSM1479436 1	24 (year)	whole blood	Type 1 Diabetes	34	Homo sapiens	white	female		whole blood	Type	EN	FT
GSM1479437 1	27 (year)	whole blood	Type 1 Diabetes	37	Homo sapiens	white	female		whole blood	Type	EN	FT
GSM1479437 1	27 (year)	whole blood	Type 1 Diabetes	37	Homo sapiens	white	female		whole blood	Type	EN	FT
GSM1479438 1	32 (year)	neutrophils	normal	44	Homo sapiens	hispanic	female		neutrophils	norm	EN	FT
GSM1479438 1	32 (year)	neutrophils	normal	44	Homo sapiens	hispanic	female		neutrophils	norm	EN	FT
		(year)	monocytes	normal		hispanic	female		monocytes	norm	EN	FT
		(year)	monocytes	normal		hispanic	female					
		(year)	B-Cells	normal		hispanic	female					
		(year)	B-Cells	normal		hispanic	female					

Unique sample name for each biological replicate

Sample attributes describing the source material

Experimental variables

Links to download data



Data mining with ArrayExpress

Data mining: Discover datasets to analyse yourself

Find RNA expression arrays from human patients with diabetes

Search results for diabetes

Filtered by organism Homo sapiens, experiment type "array assay", experiment type "rna assay"

Page 1 2 3 4 5 6 .. 11

Showing 1 - 25 of 255 experiments

Page size 25 50 100 250 500

Accession	Title	Type	Organism	Assays	Released	Processed	Raw	Views	Atlas
E-MTAB-2899	Transcriptional profiling by array of microRNAs from human peripheral blood mononuclear cells - PBMC - of Type 2 diabetes mellitus patients	microRNA profiling by array	Homo sapiens	12	01/04/2015	-		261	
E-MTAB-2896	Transcriptional Profiles in Peripheral Blood Mononuclear Cells - PBMC - from individual controls for Type 2 Diabetes Mellitus	transcription profiling by array	Homo sapiens	10	01/04/2015	-		382	
E-MTAB-2858	Transcriptional profile of PBMCs from type 1 diabetes patients treated with low-dose human interleukin-2	transcription profiling by array							
E-MTAB-2854	Transcriptional Profiles of peripheral blood mononuclear cells - PBMC - from Type 2 diabetes mellitus Patients	transcription profiling by array							

Exact match to search term

Matched EFO synonyms to search term

Matched EFO child term of search term

Export table in Tab-delimited format

Export matching



ArrayExpress – Advanced search

diabetes

Examples: E-MEXP-31, cancer, p53, Geuvadis

advanced search

Search in all experimental fields,
e.g. experiment description,
protocols, publication title...

evv:diabetes

fieldname:value

advanced search

Limit your search only to
experiments in which “diabetes” is
the value of the experimental variable

Search results for evv:diabetes

Filtered by organism **Homo sapiens**, experiment type "rna assay", experiment type "array assay"

Page 1 2 3

Showing 1 - 25 of 57 experiments

Page size 25 50 100 250 500

Accession	Title	Type	Organism	Assays	Released	Processed	Raw	Views	Atlas
E-MEXP-3907	Transcription profiling by array of human hematopoietic stem cells isolated from type 1 diabetes patients and healthy controls	transcription profiling by array	Homo sapiens	8	02/04/2017	-		341	-
E-MEXP-3904	Transcription profiling by array of human mesenchymal stromal cells from patients with type I diabetes before and after autologous hematopoietic stem cell transplantation	transcription profiling by array	Homo sapiens	12	01/12/2016	-		323	-
E-MTAB-4166	Urinary exosomal miRNA signatures in type II diabetic nephropathy patients	microRNA profiling by array	Homo sapiens	24	03/11/2016	-		368	-



ArrayExpress – Advanced search

evv:diabetes

fieldname:value

E-MEXP-3907 - Transcription profiling by array of human hematopoietic stem cells isolated from type 1 diabetes patients and healthy controls

[Display full sample-data table](#)

[Export table in Tab-delimited format](#)

8 rows

Source Name	Sample Attributes			Variables		Assay Name	Links to Data
	sex	clinical information	cell type	clinical history	clinical information		
451.1	male	from patient with type 1 diabetes	hematopoietic stem cells (CD34+)	T1D patient	from patient with type 1 diabetes	H_451.1	Download
451.2	female	from patient with type 1 diabetes	hematopoietic stem cells (CD34+)	T1D patient	from patient with type 1 diabetes	H_451.2	Download
451.3	male	from patient with type 1 diabetes	hematopoietic stem cells (CD34+)	T1D patient	from patient with type 1 diabetes	H_451.3	Download
451.4	female	from patient with type 1 diabetes	hematopoietic stem cells (CD34+)	T1D patient	from patient with type 1 diabetes	H_451.4	Download
841.1	male	healthy control	hematopoietic stem cells (CD34+)	healthy control	healthy control	H_841.1	Download
841.2	male	healthy control	hematopoietic stem cells (CD34+)	healthy control	healthy control	H_841.2	Download
841.3	male	healthy control	hematopoietic stem cells (CD34+)	healthy control	healthy control	H_841.3	Download
841.4	male	healthy control	hematopoietic stem cells (CD34+)	healthy control	healthy control	H_841.4	Download



ArrayExpress – Advanced search

diabetes

Examples: E-MEXP-31, cancer, p53, Geuvadis

advanced search

Search in all experimental fields,
e.g. experiment description,
protocols, publication title...

evv:diabetes

advanced search

fieldname:value

Limit your search only to
experiments in which “diabetes” is
the value of the experimental variable

Search results for evv:diabetes

Filtered by organism **Homo sapiens**, experiment type "rna assay", experiment type "array assay"

Page 1 2 3

Showing 1 - 25 of 57 experiments

Page size 25 50 100 250 500

Accession	Title	Type	Organism	Assays	Released	Processed	Raw	Views	Atlas
E-MEXP-3907	Transcription profiling by array of human hematopoietic stem cells isolated from type 1 diabetes patients and healthy controls	transcription profiling by array	Homo sapiens	8	02/04/2017	-		341	-
E-MEXP-3904	Transcription profiling by array of human mesenchymal stromal cells from patients with type I diabetes before and after autologous hematopoietic stem cell transplantation	transcription profiling by array	Homo sapiens	12	01/12/2016	-		323	-
E-MTAB-4166	Urinary exosomal miRNA signatures in type II diabetic nephropathy patients	microRNA profiling by array	Homo sapiens	24	03/11/2016	-		368	-



ArrayExpress – Advanced search

Field name	Search scope	Example use case
accession	Experiment primary or secondary accession	accession:E-MTAB-1234
array	Array design accession or name	array:A-AFFY-33
ev (or ef)	Experimental variable (or factor), the name of the main variable under study in an experiment. E.g. if the variable is "sex" in a human study, the researchers would be comparing between male and female samples, and "sex" is not merely an attribute the samples happen to have. Has EFO expansion .	ev:genotype
evv (or efv)	The value of an experimental variable (or factor). E.g. The values for "genotype" factor can be "wild type", "p53-/-". Has EFO expansion .	evv:"wild type"
expdesign	Experiment design type, related to the questions being addressed by the study, e.g. "time series design", "stimulus or stress design", "genetic modification design". Has EFO expansion .	expdesign:"time series"
exptype	Experiment type, related to the assay technology used. See the full list of experiment types in ArrayExpress . Has EFO expansion .	exptype:"RNA-seq of coding RNA"
gxa	Presence/absence of an ArrayExpress experiment in the Expression Atlas . Use values "true" and "false" respectively.	gxa:true
pmid	PubMed identifier for a publication.	pmid:16553887
sa	Sample attribute values. Has EFO expansion .	sa:fibroblast
sac	Sample attribute category. Find experiments that have a specific sample attribute defined, e.g. "age", "strain". Has EFO expansion .	sac:age
organism	Species of the samples. Can use common name (e.g. "mouse") or binomial nomenclature/Latin names (e.g. "Mus musculus"). Has EFO expansion .	organism:"homo sapiens"



ArrayExpress – Advanced search

Data mining: Discover datasets to analyse yourself

Find experiments comparing healthy (normal) and rheumatoid arthritis patients

evv: "rheu

rheumatic

rheumatic disease

rheumatic fever

rheumatic heart disease

Rheumatic Nodule

rheumatoid

rheumatoid arthritis

Felty's syndrome

chronic childhood arthritis

pauciarticular juvenile **rheumatoid arthritis**

systemic juvenile idiopathic arthritis

rheumatoid factor measurement

rheumatoid factor seropositivity measurement

rheumatology

Search results for evv: "rheumatoid arthritis"

Filtered by organism **Homo sapiens**

Showing 1 - 25 of 57 experiments



ArrayExpress – Advanced search

Data mining: Discover datasets to analyse yourself

Find experiments comparing healthy (normal) and rheumatoid arthritis patients

Search results for evv: "rheumatoid arthritis" AND evv: normal

Filtered by organism Homo sapiens

22 experiments

Search results for evv: "rheumatoid arthritis" AND (evv: normal OR evv:healthy)

Filtered by organism Homo sapiens

Showing 1 - 25 of 34 experiments



ArrayExpress – Advanced search

Data mining: Discover datasets to analyse yourself

*Find RNA expression arrays from human cancer samples
excluding all experiments performed on cell lines*

ArrayExpress

evv:"cancer" AND evv:"normal" NOT sac:"cell line"

Examples: E-MEXP-31, cancer, p53, Geuvadis

Home Browse Submit Help About ArrayExpress Contact Us Login

Filter search results Show more data from EMBL-EBI

Search results for evv:"cancer" AND evv:"normal" NOT sac:"cell line"

Filtered by organism Homo sapiens, experiment type "rna assay", experiment type "array assay"

Page 1 2 3 4 5 6 .. 26 Showing 1 - 25 of 626 experiments Page size 25 50 100 250 500



Let's try ArrayExpress

Browsing ArrayExpress

Hands-on activity

Discover interesting datasets related to
your research area



In pairs

www.ebi.ac.uk/arrayexpress



Let's try ArrayExpress

ArrayExpress

"Barrett's esophagus" AND ev: "disease staging" advanced search

Examples: E-MEXP-31, cancer, p53, Geuvadis

[Home](#) [Browse](#) [Submit](#) [Help](#) [About ArrayExpress](#) [Contact Us](#) [Login](#)

[Filter search results](#) [Show more data from EMBL-EBI](#)

Search results for "Barrett's esophagus" AND ev: "disease staging"

Filtered by organism [Homo sapiens](#), experiment type ["rna assay"](#), experiment type ["sequencing assay"](#)

1 experiment

Accession	Title	Type	Organism	Assays	Released	Processed	Raw	Views	Atlas
E-MTAB-4054	Whole transcriptome profiling of Esophageal adenocarcinoma and Barrett's	RNA-seq of coding RNA	Homo sapiens	63	23/07/2017	-		885	

[Export table in Tab-delimited format](#) [Export matching metadata in XML format](#) [Subscribe to RSS feed matching this search](#)



Let's try ArrayExpress

E-MTAB-4054 - Whole transcriptome profiling of Esophageal adenocarcinoma and Barrett's

Status	<i>Submitted on 11 November 2015, released on 23 July 2017, last updated yesterday</i>
Organism	Homo sapiens
Samples (51)	Click for detailed sample information and links to data
Protocols (4)	Click for detailed protocol information
Description	RNA-seq was performed on esophageal adenocarcinoma (EAC), Barrett's without dysplasia, Barrett's with low-grade dysplasia (LGD) and normal squamous esophagus tissue to find early alterations in the transcriptome level turning Barrett's dysplastic.
Experiment types	RNA-seq of coding RNA, disease state design
Contact	✉ Jesper LV Maag <j.maag@garvan.org.au>
Citation	Novel Aberrations Uncovered in Barrett's Esophagus and Esophageal Adenocarcinoma Using Whole Transcriptome Sequencing . Maag J, Fisher OM, Levert-Mignon AJ, Kaczorowski DC, Thomas ML, Hussey D, Watson D, Wettstein A, Bobryshev YV, Edwards M, Dinger ME, Lord RV, Europe PMC 28751461
MINSEQE	— — — —
	Exp. design Protocols Variables Processed Seq. reads
Files	Investigation description E-MTAB-4054.idf.txt Sample and data relationship E-MTAB-4054.sdrf.txt Click to browse all available files
Links	ENA - ERP013206 Send E-MTAB-4054 data to GENOMESPACE



Let's try ArrayExpress

E-MTAB-4054 - Whole transcriptome profiling of Esophageal adenocarcinoma and Barrett's

Display summary				Export table in Tab-delimited format		
Source Name	Protocol REF	Performer	Assay Name	Comment[technical replicate group]	Technology Type	Comment[ENA_EX]
N09	P-MTAB-47385	Dominik C. Kaczorowski	LNC_N09_R		sequencing assay	ERX1220581
N10	P-MTAB-47385	Dominik C. Kaczorowski	LNC_N10_R		sequencing assay	ERX1220582
N10	P-MTAB-47385	Dominik C. Kaczorowski	LNC_N10_R		sequencing assay	ERX1220582
B05	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B05_Human_CCGTCC_R	group1	sequencing assay	ERX1220524
B05	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B05_Human_CCGTCC_R	group1	sequencing assay	ERX1220524
B05	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B05_R	group1	sequencing assay	ERX1220525
B05	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B05_R	group1	sequencing assay	ERX1220525
LNC14	P-MTAB-47385	Dominik C. Kaczorowski	LNC_14_Human_GTGAAA_R	group10	sequencing assay	ERX1220566
LNC14	P-MTAB-47385	Dominik C. Kaczorowski	LNC_14_Human_GTGAAA_R	group10	sequencing assay	ERX1220566
				group10	sequencing assay	ERX1220567
				group10	sequencing assay	ERX1220567
				group11	sequencing assay	ERX1220575
				group11	sequencing assay	ERX1220575
				group11	sequencing assay	ERX1220576
				group11	sequencing assay	ERX1220576
				group12	sequencing assay	ERX1220579
N08	P-MTAB-47385	Dominik C. Kaczorowski	LNC_N08_Human_CCGTCC_R	group12	sequencing assay	ERX1220579
N08	P-MTAB-47385	Dominik C. Kaczorowski	LNC_N08_R	group12	sequencing assay	ERX1220580
N08	P-MTAB-47385	Dominik C. Kaczorowski	LNC_N08_R	group12	sequencing assay	ERX1220580
B06	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B06_Human_ATGTCA_R	group2	sequencing assay	ERX1220526
B06	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B06_Human_ATGTCA_R	group2	sequencing assay	ERX1220526
B06	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B06_R	group2	sequencing assay	ERX1220527
B06	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B06_R	group2	sequencing assay	ERX1220527
B07	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B07_Human_GATCAG_R	group3	sequencing assay	ERX1220528
B07	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B07_Human_GATCAG_R	group3	sequencing assay	ERX1220528

12 samples (2 normal, 5 Barrett's esophagus without dysplasia and 5 esophageal adenocarcinoma)



Let's try ArrayExpress

E-MTAB-4054 - Whole transcriptome profiling of Esophageal adenocarcinoma and Barrett's

Display summary		Showing 76 - 100 of 126 rows			Export table in Tab-delimited format	
Source Name	Protocol REF	Performer	Assay Name	Comment[technical replicate group]	Technology Type	Comment[ENA_EX]
N09	P-MTAB-47385	Dominik C. Kaczorowski	LNC_N09_R		sequencing assay	ERX1220581
N10	P-MTAB-47385	Dominik C. Kaczorowski	LNC_N10_R		sequencing assay	ERX1220582
N10	P-MTAB-47385	Dominik C. Kaczorowski	LNC_N10_R		sequencing assay	ERX1220582
B05	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B05_Human_CCGTCC_R	group1	sequencing assay	ERX1220524
B05	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B05_Human_CCGTCC_R	group1	sequencing assay	ERX1220524
B05	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B05_R	group1	sequencing assay	ERX1220525
B05	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B05_R	group1	sequencing assay	ERX1220525
LNC14	P-MTAB-47385	Do			sequencing assay	ERX1220566
LNC14	P-MTAB-47385	Do			sequencing assay	ERX1220566
LNC14	P-MTAB-47385	Do			sequencing assay	ERX1220567
LNC14	P-MTAB-47385	Dominik C. Kaczorowski	LNC_14_R	group10	sequencing assay	ERX1220567
N04	P-MTAB-47385	Dominik C. Kaczorowski	LNC_N04_Human_AGTTCC_R	group11	sequencing assay	ERX1220575
N04	P-MTAB-47385	Dominik C. Kaczorowski	LNC_N04_Human_AGTTCC_R	group11	sequencing assay	ERX1220575
N04	P-MTAB-47385	Dominik C. Kaczorowski	LNC_N04_R	group11	sequencing assay	ERX1220576
N04	P-MTAB-47385	Dominik C. Kaczorowski	LNC_N04_R	group11	sequencing assay	ERX1220576
N08	P-MTAB-47385	Dominik C. Kaczorowski	LNC_N08_Human_CCGTCC_R	group12	sequencing assay	ERX1220579
N08	P-MTAB-47385	Dominik C. Kaczorowski	LNC_N08_Human_CCGTCC_R	group12	sequencing assay	ERX1220579
N08	P-MTAB-47385	Dominik C. Kaczorowski	LNC_N08_R	group12	sequencing assay	ERX1220580
N08	P-MTAB-47385	Dominik C. Kaczorowski	LNC_N08_R	group12	sequencing assay	ERX1220580
B06	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B06_Human_ATGTCA_R	group2	sequencing assay	ERX1220526
B06	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B06_Human_ATGTCA_R	group2	sequencing assay	ERX1220526
B06	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B06_R	group2	sequencing assay	ERX1220527
B06	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B06_R	group2	sequencing assay	ERX1220527
B07	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B07_Human_GATCAG_R	group3	sequencing assay	ERX1220528
B07	P-MTAB-47385	Dominik C. Kaczorowski	LNC_B07_Human_GATCAG_R	group3	sequencing assay	ERX1220528



Let's try ArrayExpress

Disease cohort	Number of samples	Number of technical replicates
normal	17	2
Barrett's esophagus non-dysplastic	14	5
Barrett's esophagus low-grade dysplasia	8	0
esophageal adenocarcinoma	12	5
Total number of samples	51	12
Number of files (paired-end libraries)	102	24
Total number of files	126	

EMBL-EBI Resources and tools for genomics and proteomics

Discover functional genomics data
with ArrayExpress

Laura Huerta, PhD
Senior Scientific Curator
lauhuema@ebi.ac.uk
Malta, 31 July 2018

