

## The background of the slide is a teal color with a faint, repeating pattern of DNA base pairs (A, C, G, T). In the bottom right corner, the text "EMBL-EBI" is displayed in white, followed by the EMBL-EBI logo, which consists of a hexagonal grid of yellow and green dots.

eva-helpdesk@ebi.ac.uk

## Genes, genomes

European Nucleotide Archive  
1000 Genomes

Ensembl  
Ensembl Genomes

European Genome-phenome Archive  
Metagenomics portal

## Gene, protein & metabolite expression

ArrayExpress

Expression Atlas

Metabolights  
PRIDE

## Protein sequences, families & motifs

InterPro

Pfam

UniProt

## Molecular structures

Protein Data Bank in Europe  
Electron Microscopy Data Bank

## Chemical biology

ChEMBL

ChEBI

## Systems

BioModels  
Enzyme Portal

BioSamples

## Literature & ontologies

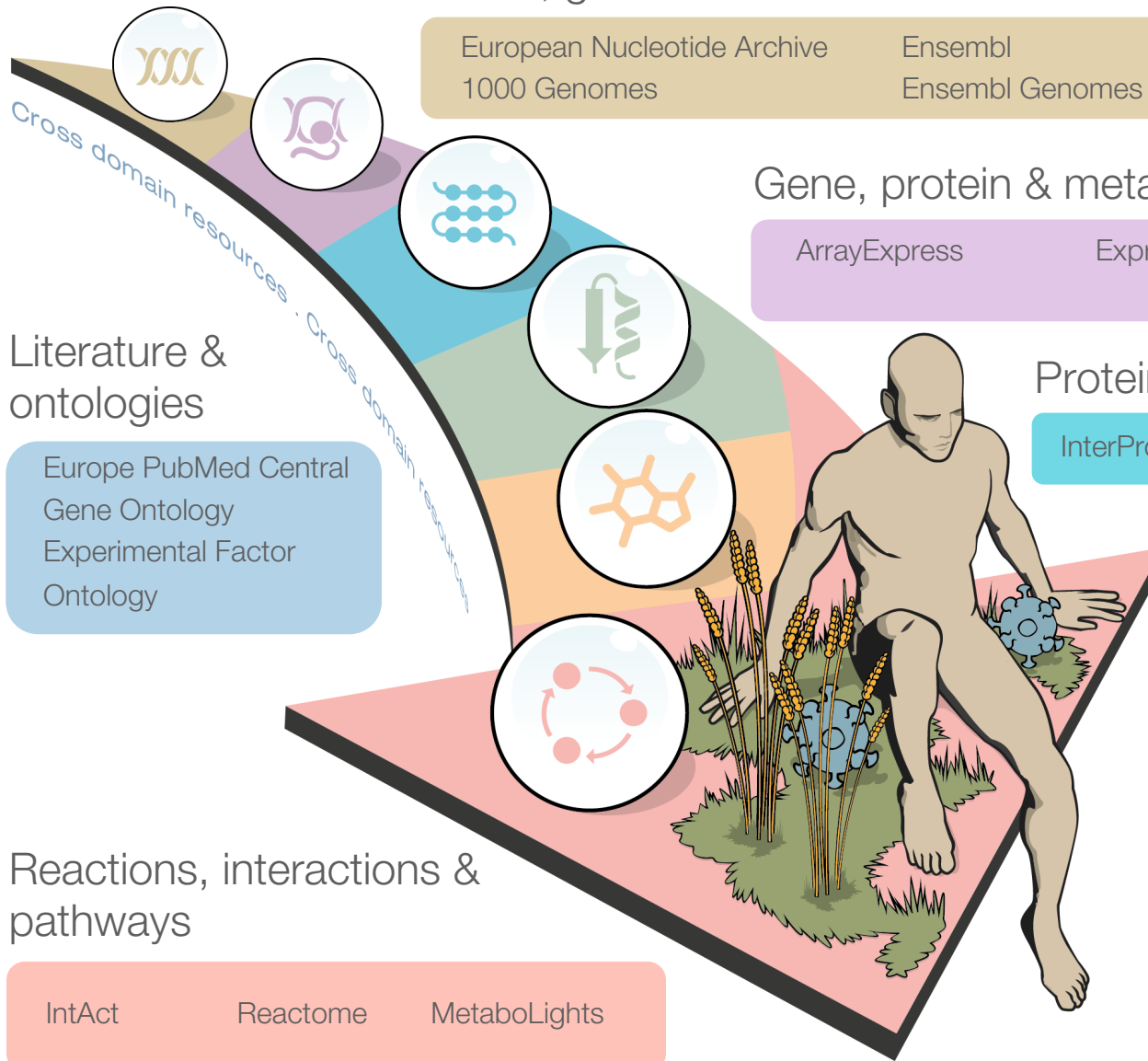
Europe PubMed Central  
Gene Ontology  
Experimental Factor  
Ontology

## Reactions, interactions & pathways

IntAct

Reactome

MetaboLights



## Genes, genomes & variation

European Nucleotide Archive  
1000 Genomes

Ensembl  
Ensembl Genomes

European Genome-phenome Archive  
Metagenomics portal

**European  
Variation  
Archive**

## Gene, protein & metabolite expression

ArrayExpress

Expression Atlas

Metabolights  
PRIDE

## Protein sequences, families & motifs

InterPro

Pfam

UniProt

## Molecular structures

Protein Data Bank in Europe  
Electron Microscopy Data Bank

## Chemical biology

ChEMBL

ChEBI

## Systems

BioModels  
Enzyme Portal

BioSamples

## Literature & ontologies

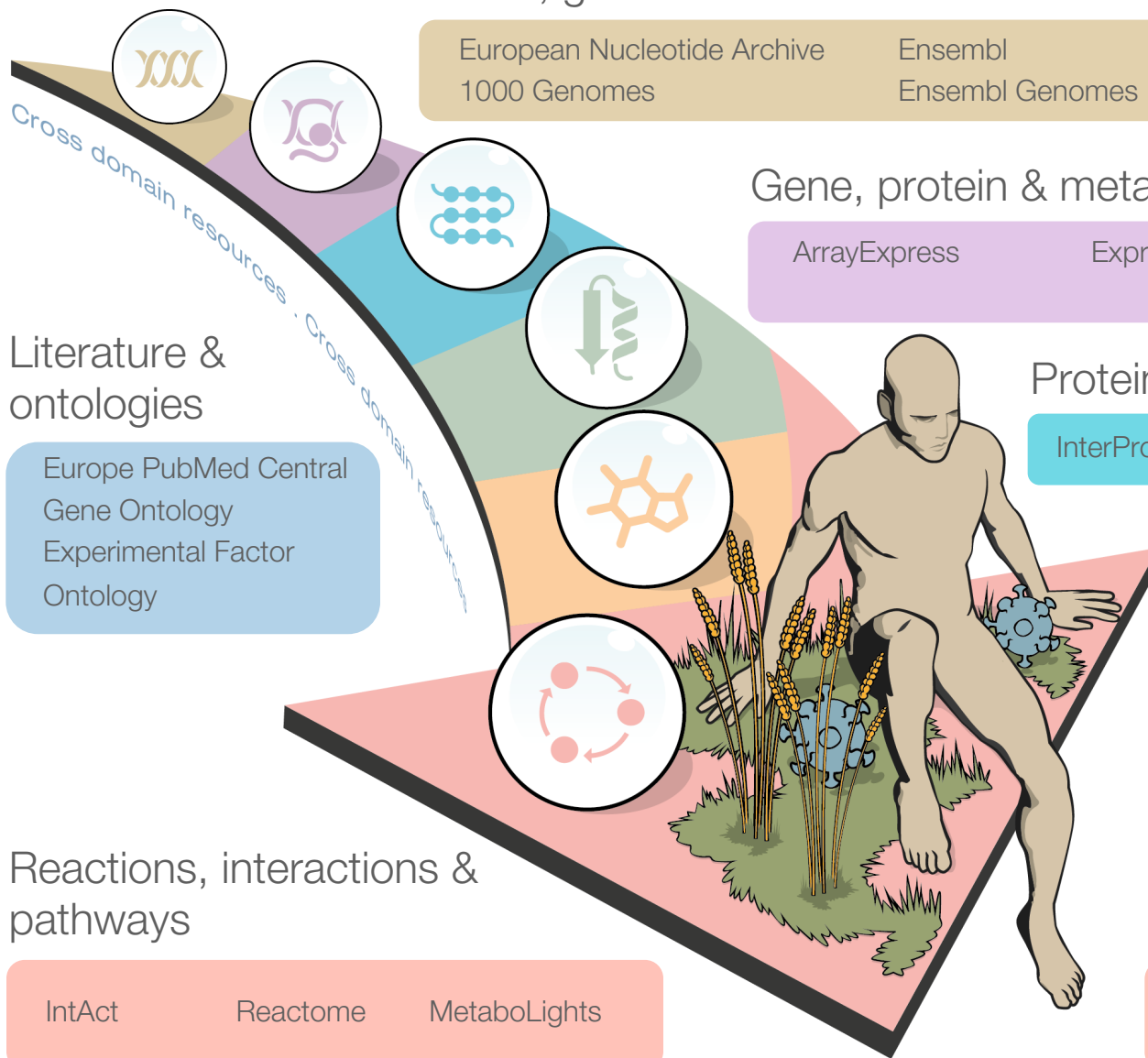
Europe PubMed Central  
Gene Ontology  
Experimental Factor  
Ontology

## Reactions, interactions & pathways

IntAct

Reactome

MetaboLights



# European Variation Archive – EVA (Eva)



- Submission based data sharing & analysis platform
- All types of variation:
- SNVs, MNVs, small indels and structural variation
- Germ line, somatic, within / cross population, potentially between species

Any variation represented as a change against a reference genome sequence

# European Variation Archive – EVA (Eva)



- Why archive VCF files?
- Administer file accessions
  - papers
  - share between researchers / labs
  - stable
- Administer variant accessions
  - TransPlant accessions
  - cluster variants based on position
  - stable

Sharing VCF files, and variant data, from a single resource

# European Variation Archive

Accept and Validate  
Submissions



Support  
dataflow from  
and to other  
EBI resources

Provide direct data  
access

Data exchange  
with peer archives



(dbVar)

Provide stable accessions

# European Variation Archive

Accept and Validate  
Submissions

Data mining

Visualization

Support  
dataflow from  
and to other  
EBI resources

Databases

Provide stable accessions



Curation

Provide direct data  
access

Ontologies

Data exchange  
with peer archives



(dbVar)



# EVA Is A Collaborative Archive

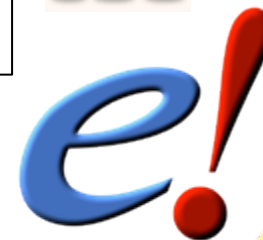


*DGVar*chive

**Direct  
Submissions**

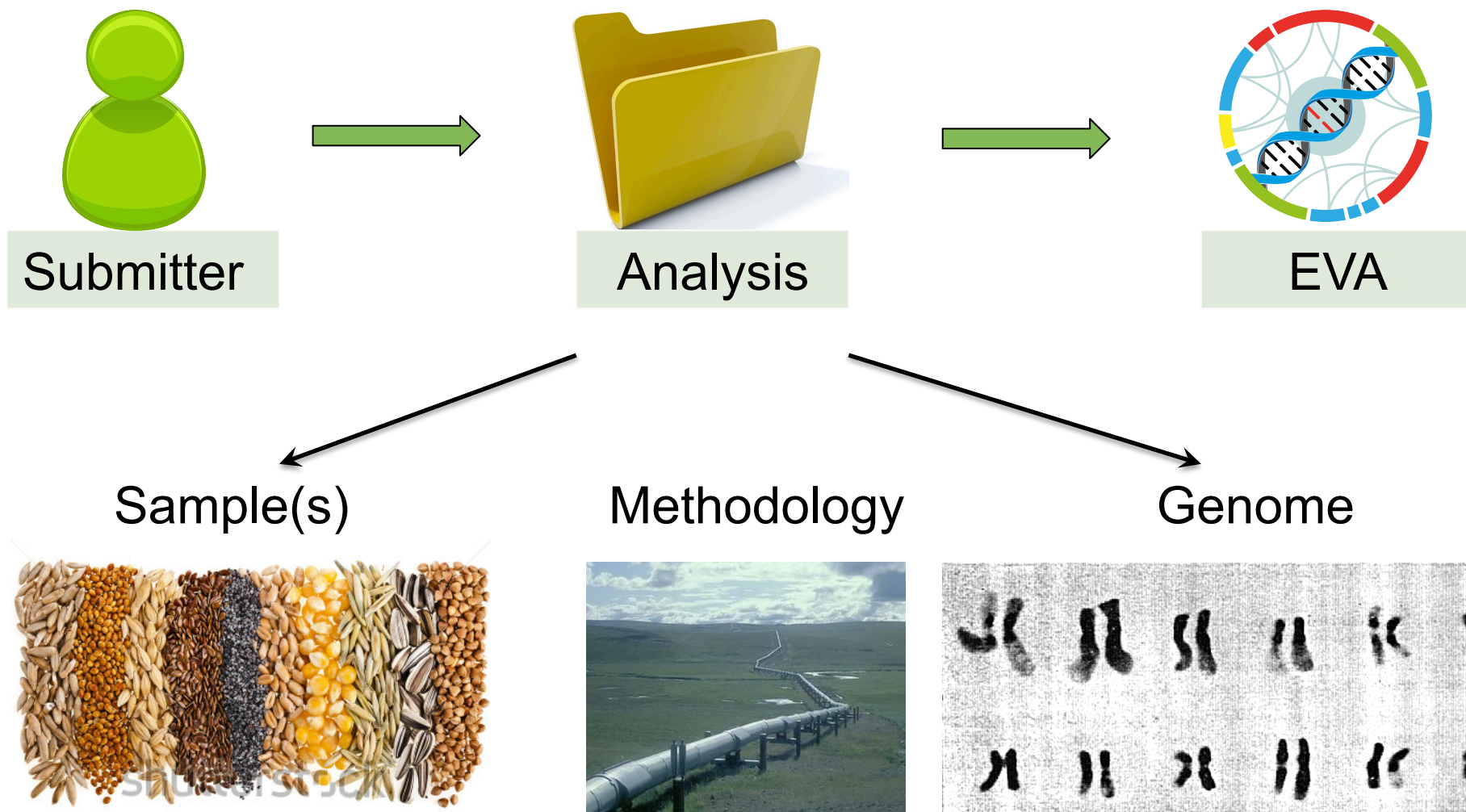


```
CTGATGGTATGGGGCCAAGAGATATATC  
AGGTACGGCTGTCATCACTTAGACCTCA  
AGGCCTGGGCATAAAGTCAGGCAGAGC  
CATGGTGCATCTGACTCCTCAGGAGAAG  
CAGGTTGGTATCAAGGTTACAGACAGG  
GGACTGACTCTCTGCTGCTATTGGTCTA
```

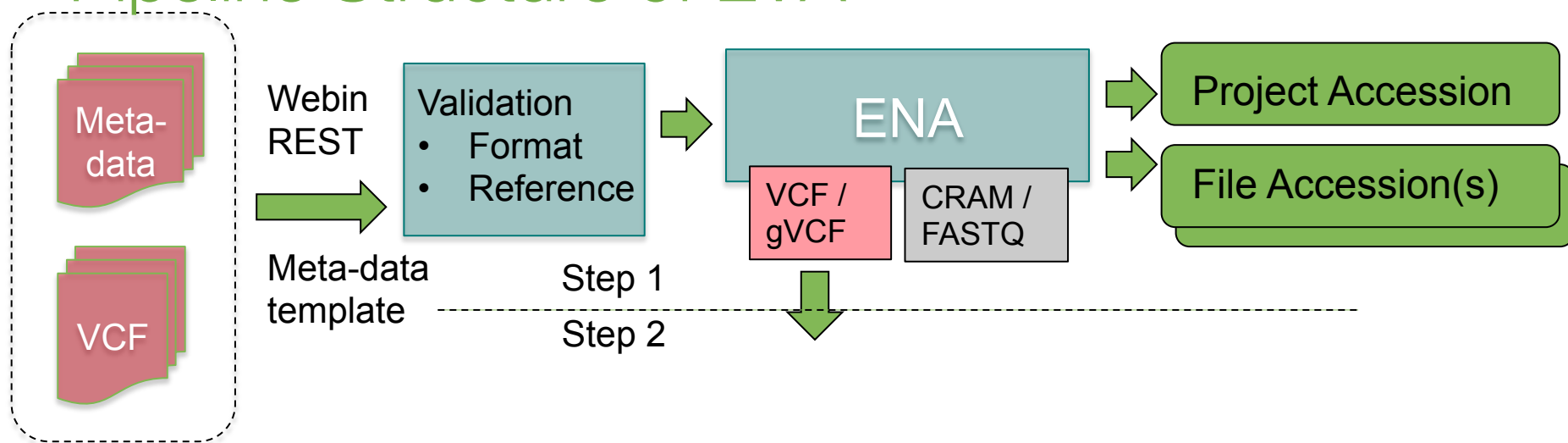




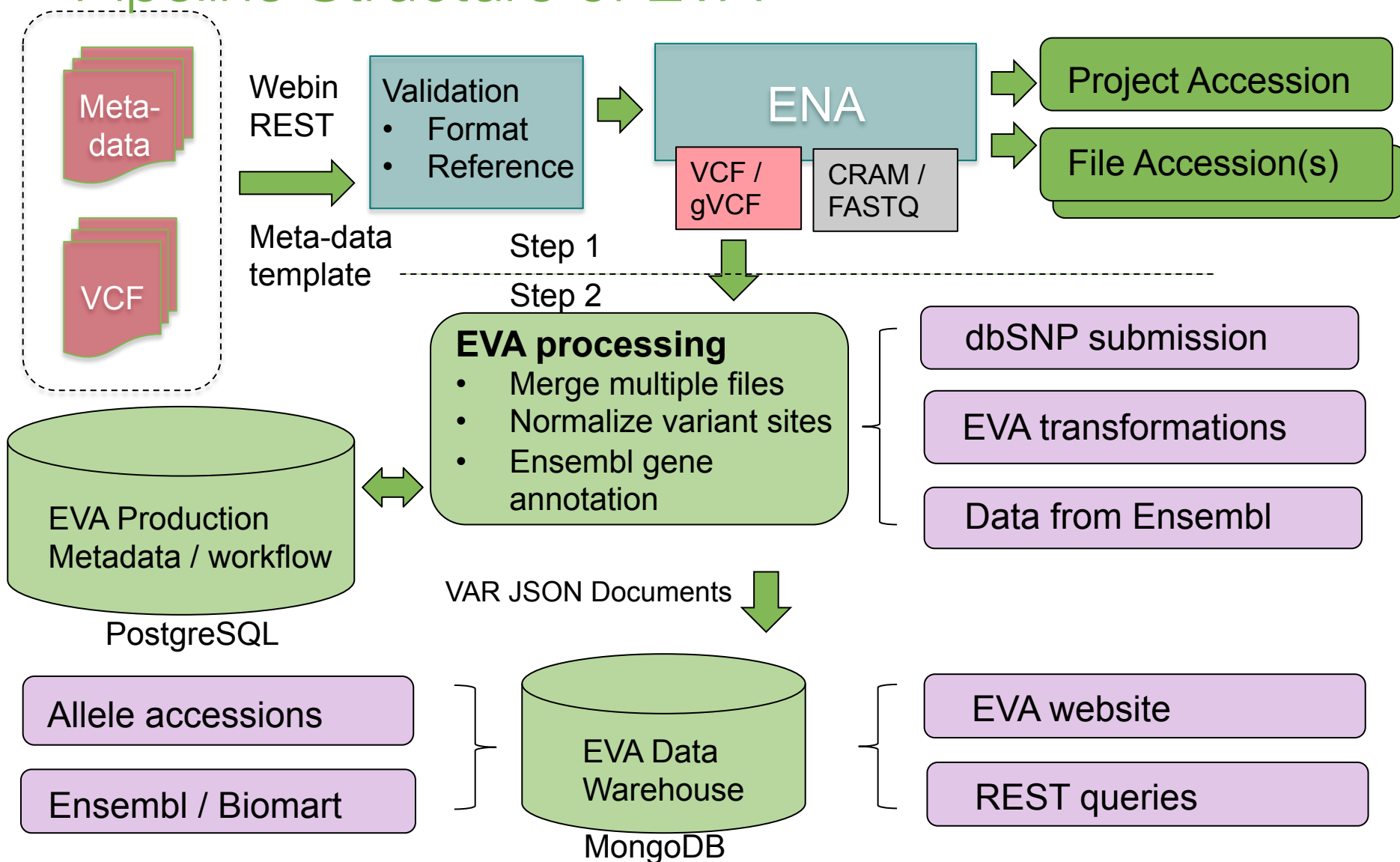
# EVA Data Model



# Pipeline Structure of EVA

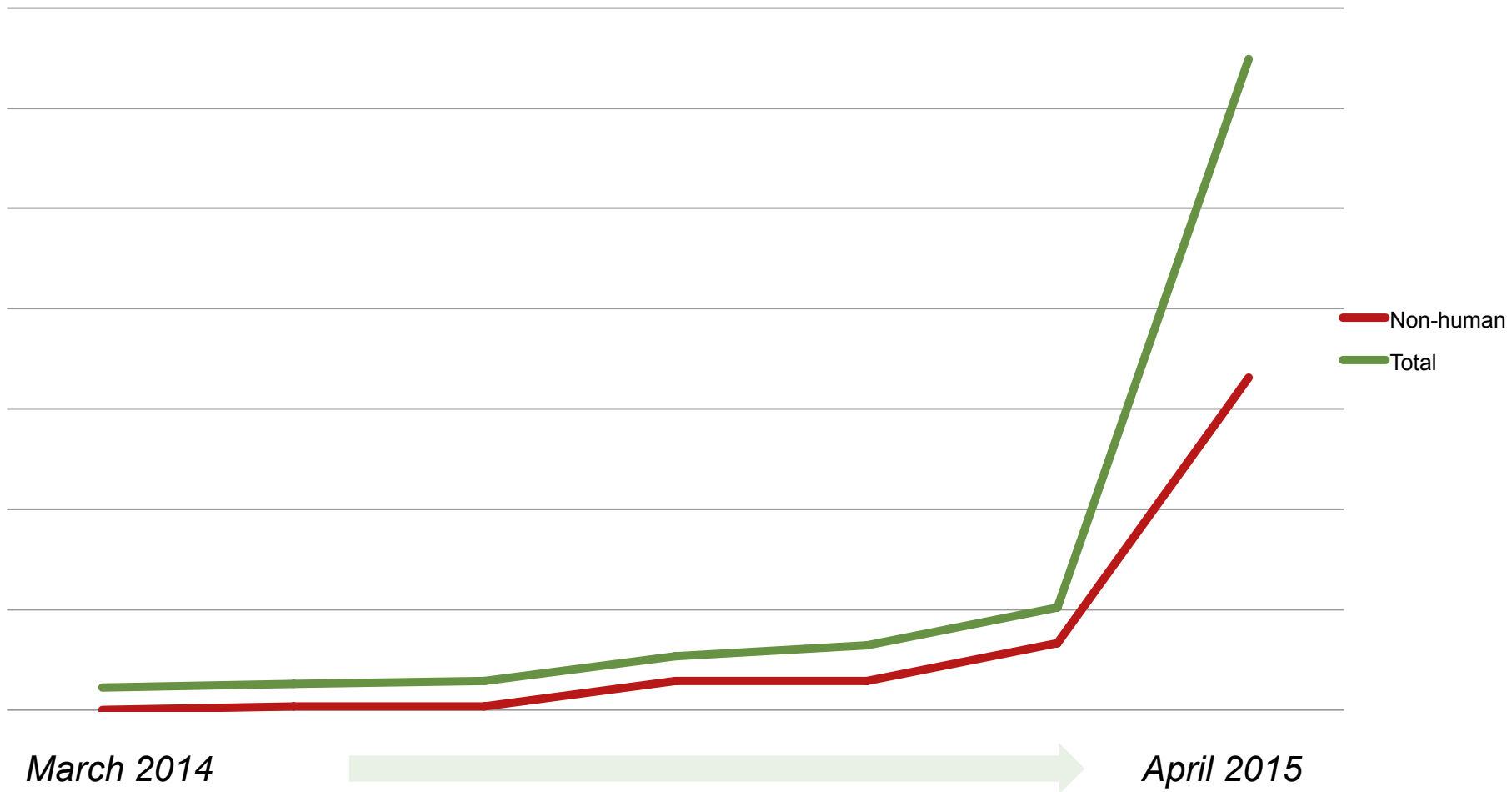


# Pipeline Structure of EVA



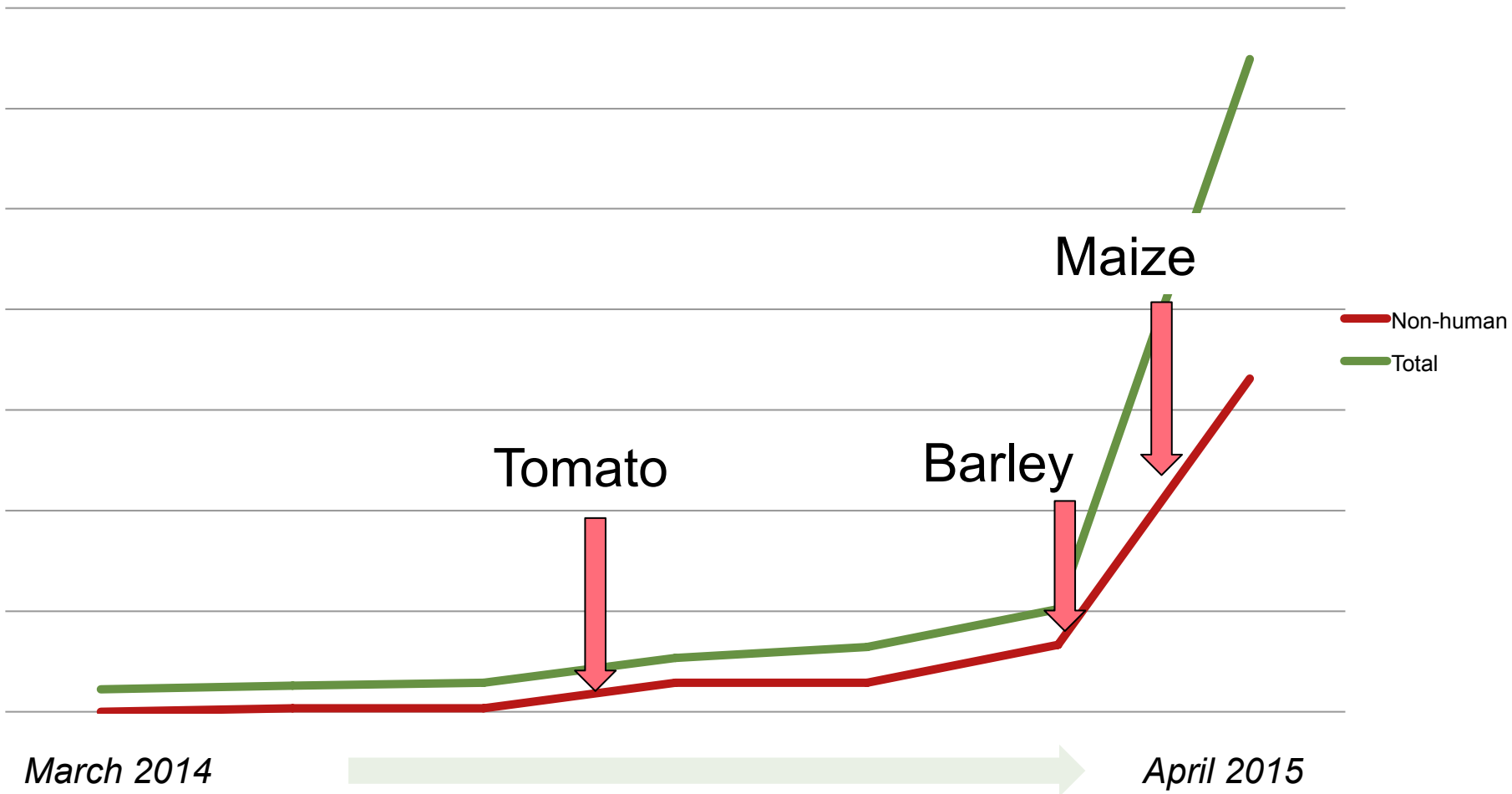
# EVA Content

- Submitted variants



# EVA Content

- Submitted variants

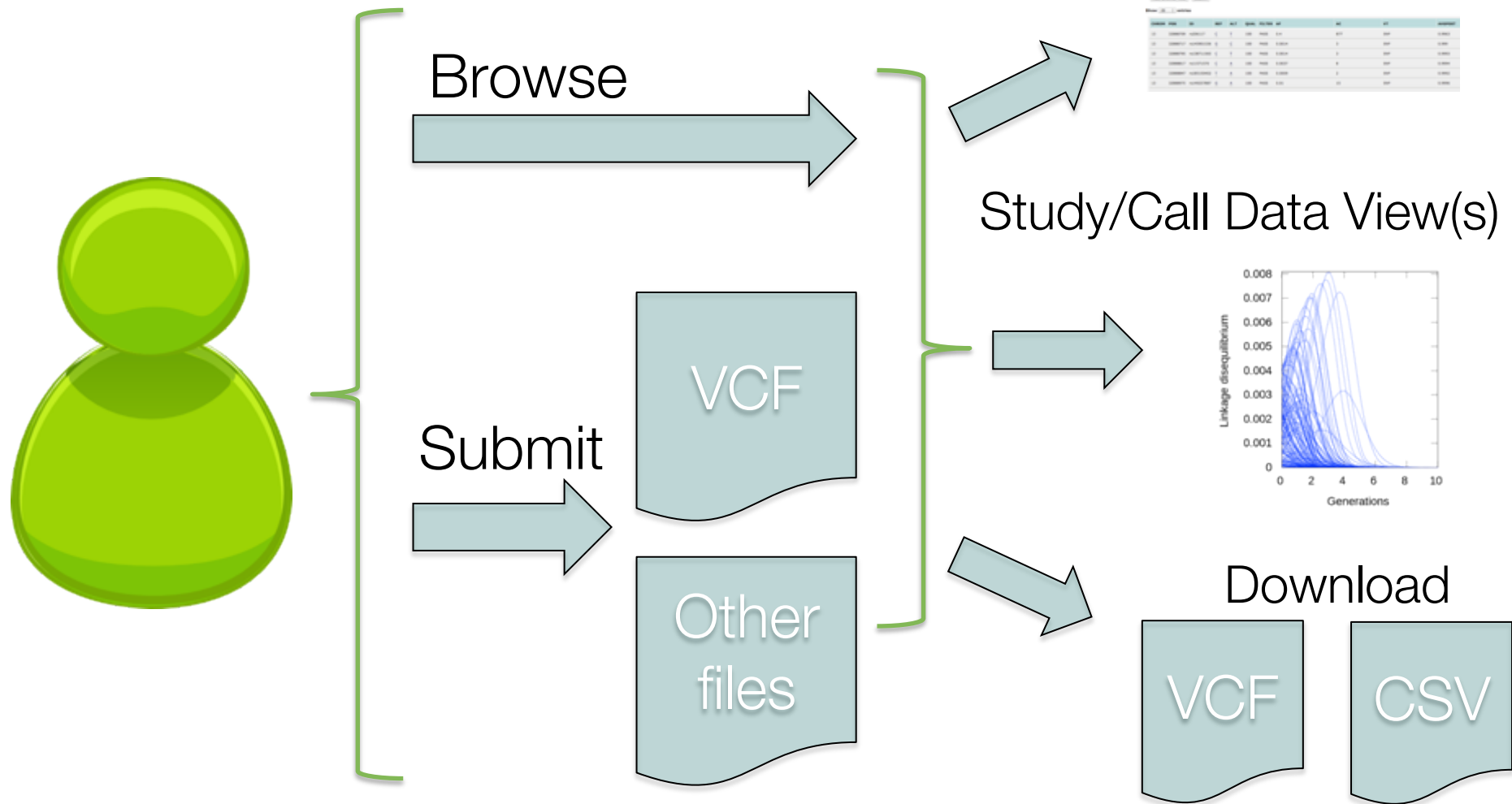


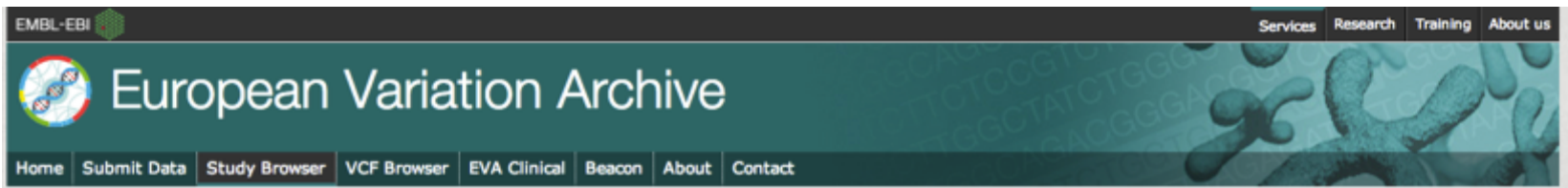
# EVA Content

- 3 main methods of increasing the data in the EVA archive:
  - 1) Large next generation sequence datasets are archived as individual studies
  - 2) User submitted datasets
  - 3) Smaller legacy datasets from dbSNP
- Shrinking percentage of the dbSNP legacy database will not be loaded as individual studies but available via dbSNP legacy track only e.g. older (array-based) technologies, very small studies



# User Experience





## Overview

The European Variation Archive is an open-access database of all types of genetic variation data from all species. The EVA provides access to highly detailed, granular, raw variant data from human, with other species to follow.

All users can [download data](#) from any study, or [submit their own data](#) to the archive. You can also query all variants in the EVA by study, gene, chromosomal location or dbSNP identifier using our [Variant Browser](#).

We will be adding new features to the EVA on a regular basis, and welcome [your comments and feedback](#).

## News

### Tweets

[Follow](#)



**Gary Saunders** @EBIvariation

8 Jan

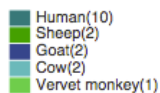
On our way to #PAGXXIII, presenting on Saturday (C01), Sunday (W370) and have a poster (P1106) on Monday. Hope to see some of you there!

[Expand](#)

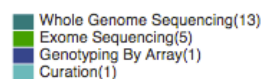
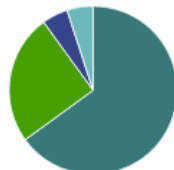
## Statistics

### Short Genetic Variations

#### Studies by Species

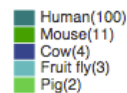
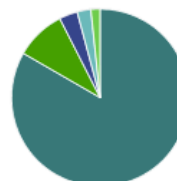


#### Studies by Type

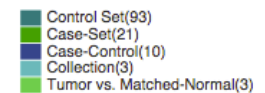


### Structural Variations

#### Studies by Species




#### Studies by Type



This web application makes an intensive use of new web technologies and standards like HTML5. Please see [About](#) for further browser compatibility notes.

# EVA Study Browser

EMBL-EBI  [Services](#) [Research](#) [Training](#) [About us](#)

## European Variation Archive

[Home](#) [Submit Data](#) [Study Browser](#) [VCF Browser](#) [EVA Clinical](#) [Beacon](#) [About](#) [Contact](#)

[Short Genetic Variations](#) [Structural Variations](#)

Short Genetic Variations Browser

Search:

Organism

- ☐ Human
- ☐ Sheep
- ☐ Goat
- ☐ Cow
- ☐ Vervet monkey
- ☐ Mosquito
- ☐ Maize
- ☐ Medaka
- ☐ Mouse
- ☐ Tomato

Type

- ☐ Whole Genome Sequencing
- ☐ Exome Sequencing
- ☐ Genotyping By Array
- ☐ Curation

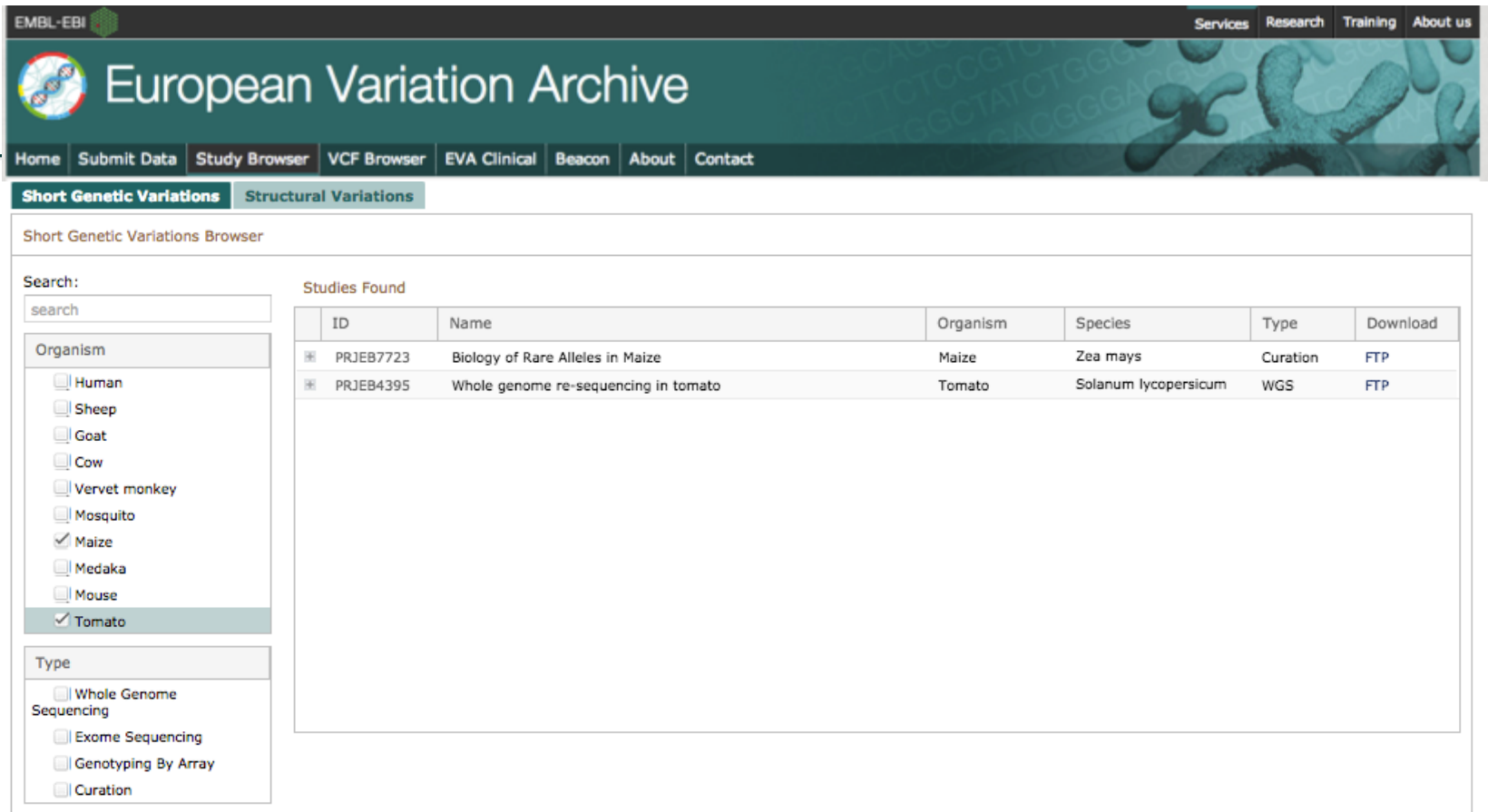
Studies Found

ID	Name	Organism	Species	Type	Download
<a href="#">PRJEB7898</a>	The Exome Aggregation Consortium (ExAC) v0.2	Human	Homo sapiens	ES	<a href="#">FTP</a>
<a href="#">PRJEB5829</a>	Genome of the Netherlands Release 5	Human	Homo sapiens	WGS	<a href="#">FTP</a>
<a href="#">PRJEB5439</a>	Exome Variant Server NHLBI Exome Sequencing Project	Human	Homo sapiens	ES	<a href="#">FTP</a>
<a href="#">PRJEB6042</a>	GEUVADIS: Genetic European Variation in Disease	Human	Homo sapiens	ES	<a href="#">FTP</a>
<a href="#">PRJEB6041</a>	UMCG Cardio GenePanel screening	Human	Homo sapiens	ES	<a href="#">FTP</a>
<a href="#">PRJEB4019</a>	1000 Genomes Phase 1 Analysis	Human	Homo sapiens	WGS, ES	<a href="#">FTP</a>
<a href="#">PRJEB7218</a>	UK10K The Department of Twin Research and Genetic Epidemiolog...	Human	Homo sapiens	WGS	<a href="#">FTP</a>
<a href="#">PRJEB7895</a>	The National FINRISK Study	Human	Homo sapiens	ES	<a href="#">FTP</a>
<a href="#">PRJEB8636</a>	deCODE Genetics Whole-Genome Sequencing Variants	Human	Homo sapiens	WGS	<a href="#">FTP</a>
<a href="#">PRJEB7217</a>	UK10K Avon Longitudinal Study of Parents and Children (ALSPAC) ...	Human	Homo sapiens	WGS	<a href="#">FTP</a>
<a href="#">PRJEB6930</a>	1000 Genomes Project phase3 release V3+	Human	Homo sapiens	WGS, ES	<a href="#">FTP</a>
<a href="#">PRJEB7923</a>	Vervet Genetic Mapping Project	Vervet monkey	Chlorocebus sabaeus	WGS	<a href="#">FTP</a>
<a href="#">PRJEB6911</a>	Sanger Institute Mouse Genomes Project v3	Mouse	Mus	WGS	<a href="#">FTP</a>
<a href="#">PRJEB6025</a>	NextGen project variation for Ovis aries	Sheep	Ovis aries	WGS	<a href="#">FTP</a>

[www.ebi.ac.uk/eva/](http://www.ebi.ac.uk/eva/)

[eva-helpdesk@ebi.ac.uk](mailto:eva-helpdesk@ebi.ac.uk)

# EVA Study Browser



The screenshot displays the EVA Study Browser interface. At the top, the EMBL-EBI logo is on the left, and navigation links for Services, Research, Training, and About us are on the right. Below this is a banner for the European Variation Archive. A secondary navigation bar includes links for Home, Submit Data, Study Browser (which is highlighted), VCF Browser, EVA Clinical, Beacon, About, and Contact. Under the Study Browser link, there are tabs for Short Genetic Variations (selected) and Structural Variations.

The main content area is titled "Short Genetic Variations Browser". It features a search section on the left with a "Search:" label and a text input field containing the word "search". Below the search field are two filter sections: "Organism" and "Type".

The "Organism" filter section lists various organisms with checkboxes: Human, Sheep, Goat, Cow, Vervet monkey, Mosquito, Maize (checked), Medaka, Mouse, and Tomato (checked). The "Type" filter section lists sequencing methods: Whole Genome Sequencing, Exome Sequencing, Genotyping By Array, and Curation.


To the right of the filters, under the heading "Studies Found", is a table with the following data:

ID	Name	Organism	Species	Type	Download
PRJEB7723	Biology of Rare Alleles in Maize	Maize	Zea mays	Curation	FTP
PRJEB4395	Whole genome re-sequencing in tomato	Tomato	Solanum lycopersicum	WGS	FTP


[www.ebi.ac.uk/eva/](http://www.ebi.ac.uk/eva/)

[eva-helpdesk@ebi.ac.uk](mailto:eva-helpdesk@ebi.ac.uk)

# EVA Study Browser – Structural Variants



ServicesResearchTrainingAbout us



## European Variation Archive

HomeSubmit DataStudy BrowserVCF BrowserEVA ClinicalBeaconAboutContact

Short Genetic VariationsStructural Variations

Structural Variations Browser

Search:

search

Organism

☐ Human  
☐ Mouse  
☐ Cow  
☐ Fruit fly  
☐ Pig  
☐ Sheep  
☐ Rhesus monkey  
☐ Zebrafish  
☐ Chimpanzee  
☐ Wolf  
☐ Horse  
☐ Sorghum  
☐ Dog

Type

☐ Control Set  
☐ Case-Set  
☐ Case-Control  
☐ Collection  
☐ Tumor vs. Matched-Normal

Studies Found

« | Page 1 of 1 | » | ↺

Studies 1 - 20 of 144

ID	Name	Organism	Species	Study Type	Download
* estd199	1000_Genomes_Consortium_...	Human	Homo sapiens	Control Set	FTP
* estd214	1000_Genomes_Consortium_...	Human	Homo sapiens	Control Set	FTP
* estd219	1000_Genomes_Consortium_...	Human	Homo sapiens	Control Set	FTP
* estd59	1000_Genomes_Consortium_...	Human	Homo sapiens	Control Set	FTP
* estd19	Ahn_et_al_2009	Human	Homo sapiens	Control Set	FTP
* nstd31	Alkan_et_al_2009	Human	Homo sapiens	Control Set	FTP
* nstd106	Alsmadi_et_al_2014	Human	Homo sapiens	Control Set	FTP
* estd195	Altshuler_et_al_2010	Human	Homo sapiens	Control Set	FTP
* nstd50	Arlt_et_al_2011	Human	Homo sapiens	Control Set	FTP
* estd176	Banerjee_et_al_2011	Human	Homo sapiens	Control Set	FTP
* estd194	Bentley_et_al_2008	Human	Homo sapiens	Control Set	FTP
* estd217	Besenbacher_et_al_2014	Human	Homo sapiens	Control Set	FTP
* nstd69	Bickhart_et_al_2012	Cow	Bos taurus	Control Set	FTP
* nstd80	Boone_et_al_2013	Human	Homo sapiens	Case Set	FTP
* nstd62	Brown_et_al_2012	Zebrafish	Danio rerio	Control Set	FTP
* nstd7	Cahan_et_al_2009	Mouse	Mus musculus	Control Set	FTP
* nstd46	Campbell_et_al_2011	Human	Homo sapiens	Control Set	FTP
* estd211	Campbell_et_al_2014	Human	Homo sapiens	Case Set	FTP
* nstd98	Campbell_et_al_2014b	Human	Homo sapiens	Case Set	FTP
* nstd13	Chen_et_al_2009	Dog	Canis lupus familiaris	Control Set	FTP

# EVA Study Browser – Structural Variants

EMBL-EBI

Services Research Training About us

## European Variation Archive

Home Submit Data Study Browser VCF Browser EVA Clinical Beacon About Contact

Short Genetic Variations **Structural Variations**

### Structural Variations Browser

Search:  
search

Organism

- ☐ Human
- ☐ Mouse
- ☐ Cow
- ☐ Fruit fly
- ☐ Pig
- ☐ Sheep
- ☐ Rhesus monkey
- ☐ Zebrafish
- ☐ Chimpanzee
- ☐ Wolf
- ☐ Horse
- ☒ Sorghum
- ☐ Dog

Type

- ☐ Control Set
- ☐ Case-Set
- ☐ Case-Control
- ☐ Collection
- ☐ Tumor vs. Matched-Normal

Studies Found

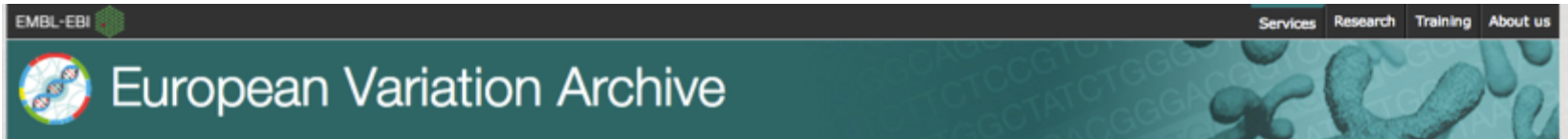
Page 1 of 1

Studies 1 - 1 of 1

ID	Name	Organism	Species	Study Type	Download
* nstd63	Zheng_et_al_2011	Sorghum	Sorghum bicolor	Control Set	FTP



# EVA 1.0



That's the data available at the current production site:  
[www.ebi.ac.uk](http://www.ebi.ac.uk)

For the remainder of this presentation, including analysis tools, I shall be discussing our upcoming first full release version


This shall be available in ca.3-4 weeks. We're in the final stages of testing

☐ Tumor vs. Matched-Normal

# EVA Study Browser – 1.0

EMBL-EBI

ServicesResearchTrainingAbout us

European Variation Archive

HomeSubmit DataStudy BrowserVCF BrowserEVA ClinicalBeaconAboutContact

Short Genetic Variations (<50bp)Structural Variations (>50bp)

Short Genetic Variations Browser

Search:  
search

Organism

- 
- ☒ Barley
- ☐ Chicken
- ☐ Cow
- ☐ Goat
- ☐ Human
- ☒ Maize
- ☐ Medaka
- ☐ Mosquito
- ☐ Mouse
- ☐ Sheep
- ☒ Tomato
- ☐ Vervet monkey

Type

- 
- ☐ Curation
- ☐ Exome Sequencing
- ☐ Genotyping By Array
- ☐ Whole Genome Sequencing


Studies Found

ID	Name	Organism	Species	Type	Download
<input checked="" type="checkbox"/> PRJEB7723	Biology of Rare Alleles in Maize	Maize	Zea mays	Curation	FTP
<input checked="" type="checkbox"/> PRJEB4395	Whole genome re-sequencing in tomato	Tomato	Solanum lycopersicum	WGS	FTP
<input checked="" type="checkbox"/> PRJEB629	Sequencing of five barley cultivars.	Barley	Hordeum vulgare subsp. v -		FTP

# EVA Study Browser – 1.0

EMBL-EBI

ServicesResearchTrainingAbout us

European Variation Archive

HomeSubmit DataStudy BrowserVCF BrowserEVA ClinicalBeaconAboutContact

Short Genetic Variations (<50bp)Structural Variations (>50bp)

Short Genetic Variations Browser

Search:  
search

Organism

- 
- ✓ Barley
- Chicken
- Cow
- Goat
- Human
- ✓ Maize
- Medaka
- Mosquito
- Mouse
- Sheep
- ✓ Tomato
- Vervet monkey

Type

- 
- Curation
- Exome Sequencing
- Genotyping By Array
- Whole Genome Sequencing

Studies Found

ID	Name	Organism	Species	Type	Download
PRJEB7723	Biology of Rare Alleles in Maize	Maize	Zea mays	Curation	FTP
PRJEB4395	Whole genome re-sequencing in tomato	Tomato	Solanum lycopersicum	WGS	FTP
PRJEB629	Sequencing of five barley cultivars.	Barley	Hordeum vulgare subsp. v -		FTP

Coming: sorghum, rice, orange, soybean, bean

Harmonizing with TransPlant


User submitted datasets...

# EVA VCF Browser

- Build view(s) of the data within EVA
- Users can query and understand data before download
- Ease of use
- Gene annotation from Ensembl
- Direct download of query results
- HTML5 technologies
- Cross-browser compatibility

# EVA VCF Browser – 1.0

EMBL-EBI

European Variation Archive

[Home](#) [Submit Data](#) [Study Browser](#) [VCF Browser](#) [EVA Clinical](#) [Beacon](#) [About](#) [Contact](#)

Services


Research


Training


About us


Filter


Clear Submit


Species 


Organism / Assembly:  
Maize / B73 RefGen\_v3 

Position 

Filter By:  
Chromosomal Location 

2:100000-400000 

Consequence Type 

search 

☐ Transcript Variant

☐ Coding Variant

☐ stop\_gained

☐ stop\_lost

☐ inframe\_insertion

☐ inframe\_deletion

☐ frameshift\_variant

☐ NMD\_transcript\_variant

☐ incomplete\_terminal\_codon...


☐ missense\_variant






☐ synonymous\_variant

☐ stop\_retained\_variant












☐ coding\_sequence\_variant


☐ feature\_elongation

Variant Browser 

  Page 2 of 7   

Variants 11 - 20 of 65


Chr	Position	Variant ID 	Alleles	Class	Most Severe Consequence Type	Protein substitution scores		View
						Polyphen2	Sift	
2	139601	vcZ2JHU51	A/C	SNV	missense_variant	-	-	 dbSNP
2	160960	vcZ2JHU52	C/A	SNV	intergenic_variant	-	-	 dbSNP
2	160979	vcZ2JHU53	G/C	SNV	intergenic_variant	-	-	 dbSNP
2	178866	vcZ2JHU54	C/T	SNV	missense_variant	-	-	 dbSNP
2	179139	vcZ2JHU55	G/A	SNV	5_prime_UTR_variant	-	-	 dbSNP
2	227241	vcZ2JHU56	G/A	SNV	intergenic_variant	-	-	 dbSNP
2	227256	vcZ2JHU57	C/T	SNV	intergenic_variant	-	-	 dbSNP
2	227270	vcZ2JHU58	G/T	SNV	intergenic_variant	-	-	 dbSNP
2	227275	vcZ2JHU59	G/A	SNV	intergenic_variant	-	-	 dbSNP
2	227297	vcZ2JHU5A	A/G	SNV	intergenic_variant	-	-	 dbSNP


Results per Page: 10  [Export as CSV](#)

Variant Data


[File and Stats](#) [Genotypes](#) [Population Stats](#) [Annotation](#)

Studies


 Biology of Rare Alleles in Maize (PRJEB7723)

Submitted\_ID 

S2\_139601

VCF data 


# EVA VCF Browser – 1.0

EMBL-EBI  [Services](#) [Research](#) [Training](#) [About us](#)

## European Variation Archive

[Home](#) [Submit Data](#) [Study Browser](#) [VCF Browser](#) [EVA Clinical](#) [Beacon](#) [About](#) [Contact](#)

### Filter

Species 

Organism / Assembly:











☐ Coding variant

- ☐ stop\_gained
- ☐ stop\_lost
- ☐ inframe\_insertion
- ☐ inframe\_deletion
- ☐ frameshift\_variant
- ☐ NMD\_transcript\_variant
- ☐ incomplete\_terminal\_codon...
- ☐ missense\_variant
- ☐ synonymous\_variant
- ☐ stop\_retained\_variant
- ☐ coding\_sequence\_variant
- ☐ feature\_elongation

### Variant Browser

Page 2 of 7

Variants 11 - 20 of 65

Chr	Position	Variant ID	Alleles	Class	Most Severe Consequence Type	Protein substitution scores		View
						Polyphen2	Sift	
2	139601	vcZ2JHU51	A/C	SNV	missense_variant	-	-	 dbSNP
2	160960	vcZ2JHU52	C/A	SNV	intergenic_variant	-	-	 dbSNP
2	160979	vcZ2JHU53	G/C	SNV	intergenic_variant	-	-	 dbSNP
2	178866	vcZ2JHU54	C/T	SNV	missense_variant	-	-	 dbSNP
2	179139	vcZ2JHU55	G/A	SNV	5_prime_UTR_variant	-	-	 dbSNP
2	227241	vcZ2JHU56	G/A	SNV	intergenic_variant	-	-	 dbSNP
2	227256	vcZ2JHU57	C/T	SNV	intergenic_variant	-	-	 dbSNP
2	227270	vcZ2JHU58	G/T	SNV	intergenic_variant	-	-	 dbSNP
2	227275	vcZ2JHU59	G/A	SNV	intergenic_variant	-	-	 dbSNP
2	227297	vcZ2JHU5A	A/G	SNV	intergenic_variant	-	-	 dbSNP

Results per Page: 10

### Variant Data

[File and Stats](#) [Genotypes](#) [Population Stats](#) [Annotation](#)

### Studies

 Biology of Rare Alleles in Maize (PRJEB7723)



VCF data



# EVA VCF Browser – 1.0

EMBL-EBI

European Variation Archive

Home Submit Data Study Browser VCF Browser EVA Clinical Beacon About Contact

Services Research Training About us

Filter

Clear Submit

Species

Organism / Assembly:  
Maize / B73 RefGen\_v3

Position

Filter By:  
Chromosomal Location  
Variant ID  
Chromosomal Location  
Ensembl Gene Symbol/Accession

Consequence Type  
search

Transcript Variant

Coding Variant

stop\_gained

stop\_lost

inframe\_insertion

inframe\_deletion

frameshift\_variant

NMD\_transcript\_variant

incomplete\_terminal\_codon...

missense\_variant

synonymous\_variant

stop\_retained\_variant

coding\_sequence\_variant

feature\_elongation

Variant Browser

Page 2 of 7

Variants 11 - 20 of 65

Chr	Position	Variant ID	Alleles	Class	Most Severe Consequence Type	Protein substitution scores		View
						Polyphen2	Sift	
2	139601	vcZ2JHU51	A/C	SNV	missense_variant	-	-	dbSNP
2	160960	vcZ2JHU52	C/A	SNV	intergenic_variant	-	-	dbSNP
2	160979	vcZ2JHU53	G/C	SNV	intergenic_variant	-	-	dbSNP
2	178866	vcZ2JHU54	C/T	SNV	missense_variant	-	-	dbSNP
2	179139	vcZ2JHU55	G/A	SNV	5_prime_UTR_variant	-	-	dbSNP
2	227241	vcZ2JHU56	G/A	SNV	intergenic_variant	-	-	dbSNP
2	227256	vcZ2JHU57	C/T	SNV	intergenic_variant	-	-	dbSNP
2	227270	vcZ2JHU58	G/T	SNV	intergenic_variant	-	-	dbSNP
2	227275	vcZ2JHU59	G/A	SNV	intergenic_variant	-	-	dbSNP
2	227297	vcZ2JHU5A	A/G	SNV	intergenic_variant	-	-	dbSNP

Results per Page: 10 Export as CSV

Variant Data

File and Stats Genotypes Population Stats Annotation

Studies

Biology of Rare Alleles in Maize (PRJEB7723)


Submitted\_ID  
S2\_139601

VCF data

# EVA VCF Browser – 1.0

EMBL-EBI

ServicesResearchTrainingAbout us

European Variation Archive

HomeSubmit DataStudy BrowserVCF BrowserEVA ClinicalBeaconAboutContact

Filter

ClearSubmit

Species

Organism / Assembly:  
Maize / B73 RefGen\_v3

Position

Filter By:  
Chromosomal Location

2:100000-400000

Consequence Type

search

☐ Transcript Variant

- ☐ Coding Variant
  - ☐ stop\_gained
  - ☐ stop\_lost
  - ☐ inframe\_insertion
  - ☐ inframe\_deletion
  - ☐ frameshift\_variant
  - ☐ NMD\_transcript\_variant
  - ☐ incomplete\_terminal\_codon...
  - ☐ missense\_variant
  - ☐ synonymous\_variant
  - ☐ stop\_retained\_variant
  - ☐ coding\_sequence\_variant
  - ☐ feature\_elongation

Variant Browser

Page 2 of 7

Variants 11 - 20 of 65

Chr	Position	Variant ID	Alleles	Class	Most Severe Consequence Type	Protein substitution scores		View
						Polyphen2	Sift	
2	139601	vcZ2JHU51	A/C	SNV	missense_variant	-	-	dbSNP
2	160960	vcZ2JHU52	C/A	SNV	intergenic_variant	-	-	dbSNP
2	160979	vcZ2JHU53	G/C	SNV	intergenic_variant	-	-	dbSNP
2	178866	vcZ2JHU54	C/T	SNV	missense_variant	-	-	dbSNP
2	179139	vcZ2JHU55	G/A	SNV	5_prime_UTR_variant	-	-	dbSNP
2	227241	vcZ2JHU56	G/A	SNV	intergenic_variant	-	-	dbSNP
2	227256	vcZ2JHU57	C/T	SNV	intergenic_variant	-	-	dbSNP
2	227270	vcZ2JHU58	G/T	SNV	intergenic_variant	-	-	dbSNP
2	227275	vcZ2JHU59	G/A	SNV	intergenic_variant	-	-	dbSNP
2	227297	vcZ2JHU5A	A/G	SNV	intergenic_variant	-	-	dbSNP

Results per Page: 10Export as CSV

Variant Data

File and StatsGenotypesPopulation StatsAnnotation

Studies

Biology of Rare Alleles in Maize (PRJEB7723)

Submitted\_ID  
S2\_139601

VCF data

# Ensembl Plants VEP

[BLAST](#) | [BioMart](#) | [Tools](#) | [Downloads](#) | [Help](#) | [Documentation](#)

Search Ensembl Plants...

Login/Register

**Tools**  
 We provide  
 In the near f  
 Currently av  
 Name  
 Variant Eff  
  
 BLAST/BL  
 BioMart  
 Assembly  
 ID History  
 Ensembl G  
 Ensembl G  
 Ensembl Plan

**Personal Data**  
[Login](#)  
[Register](#)  
[Lost Password](#)  

Custom Data

- Add your data
- Attach DAS
- Manage Data
- Features on Karyotype

Manage Configurations

- Configurations for this page
- All configurations
- Configuration sets

Online Tools

- Variant Effect Predictor**
- Assembly Converter
- ID History Converter
- Data Slicer

[Help](#)

### Variant Effect Predictor:

This tool takes a list of variant positions and alleles, and predicts the effects of each of these on overlapping transcripts and regulatory regions annotated in Ensembl. The tool accepts substitutions, insertions and deletions as input, see [data formats](#).

Upload is limited to 750 variants; lines after the limit will be ignored. Users with more than 750 variations can split files into smaller chunks, use the standalone [perl script](#) or the [variation API](#). See also [full documentation](#)

NB: Ensembl now by default uses Sequence Ontology terms to describe variation consequences. See [this page](#) for details

#### Input file

Species:

Name for this data (optional):

Paste data:

Upload file:  No file chosen

or provide file URL:

Input file format:

#### Options

Transcript database to use:

☒ Ensembl transcripts

☐ RefSeq and other transcripts

Type of consequences to display:

Check for existing co-located variants:

Return results for variants in coding regions only: ☐

Show Ensembl protein identifiers where available: ☐


Show HGVS identifiers for variants where available: ☐

[Cookies](#) | [Help](#)

# EVA VCF Browser – 1.0

EMBL-EBI

ServicesResearchTrainingAbout us

European Variation Archive

HomeSubmit DataStudy BrowserVCF BrowserEVA ClinicalBeaconAboutContact

Filter

ClearSubmit

Species

Organism / Assembly:  
Maize / B73 RefGen\_v3

Position

Filter By:  
Chromosomal Location

2:100000-400000

Consequence Type

search

☐ Transcript Variant

- ☐ Coding Variant
  - ☐ stop\_gained
  - ☐ stop\_lost
  - ☐ inframe\_insertion
  - ☐ inframe\_deletion
  - ☐ frameshift\_variant
  - ☐ NMD\_transcript\_variant
  - ☐ incomplete\_terminal\_codon...
  - ☐ missense\_variant
  - ☐ synonymous\_variant
  - ☐ stop\_retained\_variant
  - ☐ coding\_sequence\_variant
  - ☐ feature\_elongation

Variant Browser

Page 2 of 7

Variants 11 - 20 of 65

Chr	Position	Variant ID	Alleles	Class	Most Severe Consequence Type	Protein substitution scores		View
						Polyphen2	Sift	
2	139601	vcZ2JHU51	A/C	SNV	missense_variant	-	-	dbSNP
2	160960	vcZ2JHU52	C/A	SNV	intergenic_variant	-	-	dbSNP
2	160979	vcZ2JHU53	G/C	SNV	intergenic_variant	-	-	dbSNP
2	178866	vcZ2JHU54	C/T	SNV	missense_variant	-	-	dbSNP
2	179139	vcZ2JHU55	G/A	SNV	5_prime_UTR_variant	-	-	dbSNP
2	227241	vcZ2JHU56	G/A	SNV	intergenic_variant	-	-	dbSNP
2	227256	vcZ2JHU57	C/T	SNV	intergenic_variant	-	-	dbSNP
2	227270	vcZ2JHU58	G/T	SNV	intergenic_variant	-	-	dbSNP
2	227275	vcZ2JHU59	G/A	SNV	intergenic_variant	-	-	dbSNP
2	227297	vcZ2JHU5A	A/G	SNV	intergenic_variant	-	-	dbSNP

Results per Page: 10Export as CSV

Variant Data

File and StatsGenotypesPopulation StatsAnnotation

Studies

Biology of Rare Alleles in Maize (PRJEB7723)


Submitted\_ID  
S2\_139601

VCF data

# EVA VCF Browser – 1.0

EMBL-EBI

ServicesResearchTrainingAbout us

European Variation Archive

HomeSubmit DataStudy BrowserVCF BrowserEVA ClinicalBeaconAboutContact

Filter

Variant Browser

ClearSubmit

Species

Page 2 of 7

Variants 11 - 20 of 65

Chr	Position	Variant ID	Alleles	Class	Most Severe Consequence Type	Protein substitution scores		View
						Polyphen2	Sift	

```
##fileformat=VCFv4.0
##fileDate=20140213
##source=AllZeaGBSv2.7_publicSamples_imputedV3b
##reference=B73RefGenV3
##phasing=no
##INFO=<ID=MQ,Number=.,Type=Float,Description="RMS mapping quality">
##INFO=<ID=Submitted_ID,Number=1,Type=String,Description="Submitter ID">
##FORMAT=<ID=GT,Number=1,Type=String,Description="Genotype">
##contig=<ID=1,species="Zea mays",accession=GK000031.3>
##contig=<ID=2,species="Zea mays",accession=GK000032.3>
##contig=<ID=3,species="Zea mays",accession=GK000033.3>
##contig=<ID=4,species="Zea mays",accession=CM000780.3>
##contig=<ID=5,species="Zea mays",accession=CM000781.3>
##contig=<ID=6,species="Zea mays",accession=CM000782.3>
##contig=<ID=7,species="Zea mays",accession=GK000034.3>
##contig=<ID=8,species="Zea mays",accession=CM000784.3>
##contig=<ID=9,species="Zea mays",accession=CM000785.3>
##contig=<ID=10,species="Zea mays",accession=CM000786.3>
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT SAMEA2833119 SAMEA2830877 SAMEA2837965 SAMEA2839535 SAMEA2839706 SAMEA2843832 SAMEA2829711 SAMEA2828547 SAMEA2844507
2 139601 vcZ2JHU51 A C . . Submitted_ID=S2_139601 GT 0/0 0/0 0/0 ./. 0/0 0/0 ./. 0/0 0/0 0/0 ./. ./. ./. ./. ./. ./. ./. 0/0 0/0 0/0 0/0 0/0 0/0 0/0 0/0 0/0 0/0 ./.

```

Studies

☐ frameshift\_variant

☐ NMD\_transcript\_variant

☐ incomplete\_terminal\_codon...

☐ missense\_variant

☐ synonymous\_variant

☐ stop\_retained\_variant

☐ coding\_sequence\_variant

☐ feature\_elongation

Biology of Rare Alleles in Maize (PRJEB7723)

Submitted\_ID ⓘ


S2\_139601

VCF data +

# EVA VCF Browser – 1.0

EMBL-EBI

Services Research Training About us



# European Variation Archive

Home Submit Data Study Browser VCF Browser EVA Clinical Beacon About Contact

## Filter

Species

Organism / Assembly:

Maize / B73 RefGen\_v3

Position

Filter By:

Chromosomal Location

2:100000-400000

Consequence Type

search

☐ Transcript Variant

- ☐ Coding Variant
  - ☐ stop\_gained
  - ☐ stop\_lost
  - ☐ inframe\_insertion
  - ☐ inframe\_deletion
  - ☐ frameshift\_variant
  - ☐ NMD\_transcript\_variant
  - ☐ incomplete\_terminal\_codon...
  - ☐ missense\_variant
  - ☐ synonymous\_variant
  - ☐ stop\_retained\_variant
  - ☐ coding\_sequence\_variant
  - ☐ feature\_elongation

Protein Substitution Score

Polyphen2 >: ex: 0.5

Sift <: ex: 0.1

## Variant Browser

« « Page 2 of 7 » » ↺

Variants 11 - 20 of 65

Chr	Position	Variant ID	Alleles	Class	Most Severe Consequence Type	Protein substitution scores		View
						Polyphen2	Sift	
2	139601	vc22JHU51	A/C	SNV	missense_variant	-	-	dbSNP
2	160960	vc22JHU52	C/A	SNV	intergenic_variant	-	-	dbSNP
2	160979	vc22JHU53	G/C	SNV	intergenic_variant	-	-	dbSNP
2	178866	vc22JHU54	C/T	SNV	missense_variant	-	-	dbSNP
2	179139	vc22JHU55	G/A	SNV	5_prime_UTR_variant	-	-	dbSNP
2	227241	vc22JHU56	G/A	SNV	intergenic_variant	-	-	dbSNP
2	227256	vc22JHU57	C/T	SNV	intergenic_variant	-	-	dbSNP
2	227270	vc22JHU58	G/T	SNV	intergenic_variant	-	-	dbSNP
2	227275	vc22JHU59	G/A	SNV	intergenic_variant	-	-	dbSNP
2	227297	vc22JHU5A	A/G	SNV	intergenic_variant	-	-	dbSNP

Results per Page: 10

## Variant Data

File and Stats

Genotypes

Population Stats


Annotation

## Genotypes

Study		Samples Count
Biology of Rare Alleles in Maize (PRJEB7723)		17280
Sample	Genotype	
SAMEA2830759	./.	
SAMEA2830758	0/0	
SAMEA2830757	./.	
SAMEA2830756	0/0	
SAMEA2830755	./.	
SAMEA2830754	./.	



# EVA VCF Browser – 1.0

EMBL-EBI  [Services](#) [Research](#) [Training](#) [About us](#)


## European Variation Archive

[Home](#) [Submit Data](#) [Study Browser](#) [VCF Browser](#) [EVA Clinical](#) [Beacon](#) [About](#) [Contact](#)

### Filter


Species 


Organism / Assembly:


Maize / B73 RefGen\_v3 

Position 

Filter By:

Chromosomal Location 

2:100000-400000 

Consequence Type 

search 

☐ Transcript Variant

☐ Coding Variant

☐ stop\_gained

☐ stop\_lost

☐ inframe\_insertion

☐ inframe\_deletion

☐ frameshift\_variant

☐ NMD\_transcript\_variant

☐ incomplete\_terminal\_codon...

☐ missense\_variant




☐ synonymous\_variant

☐ stop\_retained\_variant











☐ coding\_sequence\_variant

☐ feature\_elongation

### Variant Browser

Page 2 of 7   

Variants 11 - 20 of 65

Chr	Position	Variant ID	Alleles	Class	Most Severe Consequence Type	Protein substitution scores		View
						Polyphen2	Sift	
2	139601	vc2ZJHU51	A/C	SNV	missense_variant	-	-	 dbSNP
2	160960	vc2ZJHU52	C/A	SNV	intergenic_variant	-	-	 dbSNP
2	160979	vc2ZJHU53	G/C	SNV	intergenic_variant	-	-	 dbSNP
2	178866	vc2ZJHU54	C/T	SNV	missense_variant	-	-	 dbSNP
2	179139	vc2ZJHU55	G/A	SNV	5_prime_UTR_variant	-	-	 dbSNP
2	227241	vc2ZJHU56	G/A	SNV	intergenic_variant	-	-	 dbSNP
2	227256	vc2ZJHU57	C/T	SNV	intergenic_variant	-	-	 dbSNP
2	227270	vc2ZJHU58	G/T	SNV	intergenic_variant	-	-	 dbSNP
2	227275	vc2ZJHU59	G/A	SNV	intergenic_variant	-	-	 dbSNP
2	227297	vc2ZJHU5A	A/G	SNV	intergenic_variant	-	-	 dbSNP




Results per Page: 10

[Export as CSV](#)

### Variant Data

[File and Stats](#) [Genotypes](#) [Population Stats](#) [Annotation](#)

### Annotations

Page 1 of 1   

Transcripts 1 - 1 of 1

Ensembl Gene ID	Ensembl Gene Symbol	Ensembl Transcript ID	SO Term(s)	Biotype	Codon	cDna Position	AA Change
GRMZM2G046590		GRMZM2G046590...	missense_variant	protein_coding	gaA/g...	1517	E/D

# EVA VCF Browser – 1.0

EMBL-EBI

ServicesResearchTrainingAbout us

EnsemblPlants

BLAST | BioMart | Tools | Downloads | Help | Documentation

Zea mays

Location: 2:134,843-139,880

Gene: GRMZM2G046590

Trans: GRMZM2G046590\_T01

Login/Register

Search Ensembl Plants...

Gene-based displays

Summary

Splice variants

Transcript comparison

Supporting evidence

Gene alleles

Sequence

Secondary Structure

External references

Regulation

Literature

Ontology

GO: biological process

GO: cellular component

Plant Compara

Genomic alignments

Gene tree

Gene gain/loss tree

Orthologues

Paralogues

Pan-taxonomic Compara

Gene Tree

Orthologues

Phenotype

Genetic Variation

Variation table

Structural variation

Variation image

External data

Gene expression

Personal annotation

ID History

Gene history

Configure this page

Add your data

Export data

Share this page

Bookmark this page

Ensembl Plants is produced in collaboration with Gramene

Gene: GRMZM2G046590

LocationChromosome 2: 134,843-139,880 forward strand.

About this geneThis gene has 1 transcript ([splice variant](#)), [64 orthologues](#) and [23 paralogues](#).

Transcripts

Show transcript table

Summary ⓘ

Gene typeProtein coding

Annotation MethodGene annotation by [Gramene](#) through an automated, evidence-based method

Go to Region in Detail for more tracks and navigation options (e.g. zooming)

GRMZM2G046590

25.04 kb

128kb128kb130kb132kb134kb136kb138kb140kb142kb144kb146kb148kb

Forward strand

Reverse strand

Protein Coding

protein coding

AC185411.3-Contig150

AC205258.3-Contig33

Configuring the display

Tip: use the "Configure this page" link on the left to show additional data in this region.

Ensembl Plants release 27 - June 2015 © EBI

About Ensembl Genomes | Contact Us | EMBL-EBI Terms of use | Privacy | Cookies | Help

Ensembl Gene ID	Ensembl Gene Symbol	Ensembl Transcript ID	SO Term(s)	Biotype	Codon	cDNA Position	AA Change
GRMZM2G046590		GRMZM2G046590...	missense_variant	protein_coding	gaA/g...	1517	E/D

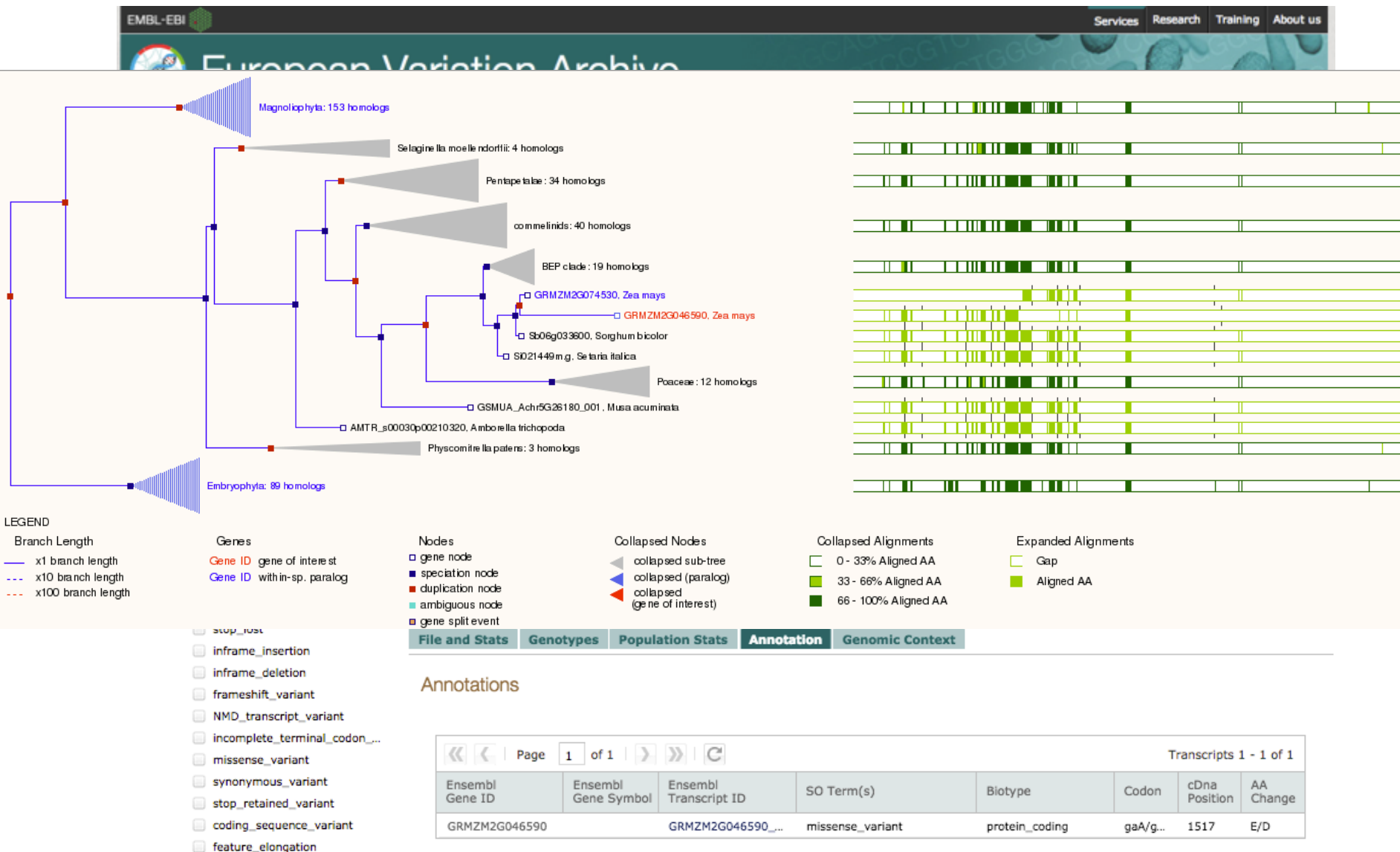
☐ synonymous\_variant

☐ stop\_retained\_variant

☐ coding\_sequence\_variant

☐ feature\_elongation


# EVA VCF Browser – 1.0



# EVA VCF Browser – 1.0

EMBL-EBI

Services Research Training About us

 European Variation Archive

Home Submit Data Study Browser VCF Browser EVA Clinical Beacon About Contact

## Filter

Clear Submit

Species

Organism / Assembly:

Maize / B73 RefGen\_v3

Position

Filter By:

Chromosomal Location

2:100000-400000

Consequence Type











search

- ☐ Transcript Variant
  - ☐ Coding Variant
    - ☐ stop\_gained
    - ☐ stop\_lost
    - ☐ inframe\_insertion
    - ☐ inframe\_deletion
    - ☐ frameshift\_variant
    - ☐ NMD\_transcript\_variant
    - ☐ incomplete\_terminal\_codon...
    - ☐ missense\_variant
    - ☐ synonymous\_variant
    - ☐ stop\_retained\_variant
    - ☐ coding\_sequence\_variant
    - ☐ feature\_elongation

## Variant Browser

Page 2 of 7

Variants 11 - 20 of 65

Chr	Position	Variant ID	Alleles	Class	Most Severe Consequence Type	Protein substitution scores		View
						Polyphen2	Sift	
2	139601	vcZ2JHU51	A/C	SNV	missense_variant	-	-	 dbSNP
2	160960	vcZ2JHU52	C/A	SNV	intergenic_variant	-	-	 dbSNP
2	160979	vcZ2JHU53	G/C	SNV	intergenic_variant	-	-	 dbSNP
2	178866	vcZ2JHU54	C/T	SNV	missense_variant	-	-	 dbSNP
2	179139	vcZ2JHU55	G/A	SNV	5_prime_UTR_variant	-	-	 dbSNP
2	227241	vcZ2JHU56	G/A	SNV	intergenic_variant	-	-	 dbSNP
2	227256	vcZ2JHU57	C/T	SNV	intergenic_variant	-	-	 dbSNP
2	227270	vcZ2JHU58	G/T	SNV	intergenic_variant	-	-	 dbSNP
2	227275	vcZ2JHU59	G/A	SNV	intergenic_variant	-	-	 dbSNP
2	227297	vcZ2JHU5A	A/G	SNV	intergenic_variant	-	-	 dbSNP

Results per Page: 10 Export as CSV

## Variant Data

File and Stats Genotypes Population Stats Annotation

## Annotations

Page 1 of 1

Transcripts 1 - 1 of 1

Ensembl Gene ID	Ensembl Gene Symbol	Ensembl Transcript ID	SO Term(s)	Biotype	Codon	cDna Position	AA Change
GRMZM2G046590		GRMZM2G046590...	missense_variant	protein_coding	gaA/g...	1517	E/D

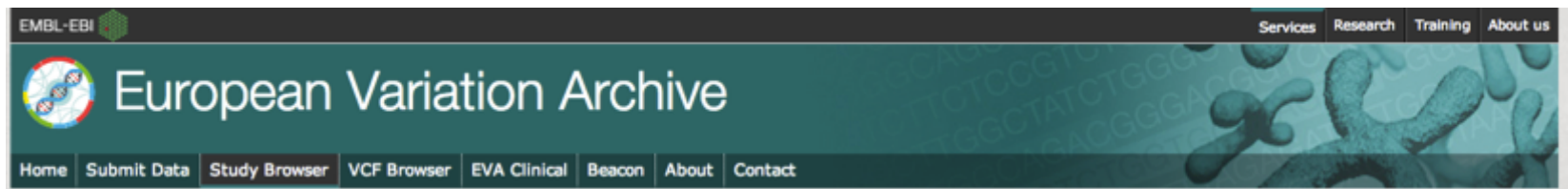
# EVA API

- EVA VCF browser is our GUI view of web service results
- Efficient programmatic access through a RESTful web services API
- All EVA data available regardless of the programming language
- Results provided as JSON objects: easily parsed by Python, R, JAVA, for example
- Web services for:
  - files, segments, studies, variants
  - full documentation at EVA website (1.0)

# EVA needs your input!

- EVA is now in beta-release, please suggest changes to make it more useful
- EVA 1.0, first full release, July 2015, comments and suggestions are the foundations of changes
- Data submissions are key, ideally with genotypes
- Contact at [eva-helpdesk@ebi.ac.uk](mailto:eva-helpdesk@ebi.ac.uk)
  - I'm also around this week

# Submit data to EVA



## European Variation Archive submissions

EVA follows the infrastructure of fellow EMBL-EBI resources European Nucleotide Archive (ENA) and European Genome-phenome Archive (EGA) to accept, archive, and accession [VCF files](#). Submissions consist of VCF file(s) and metadata that describe sample(s), experiment(s), and analysis that produced the variant and/or genotype call(s).

EVA works in collaboration with the Database of Genomics Variants Archive (DGVA) to accession and archive structural variants. DGVA relies on a template based submission process that is explained in detail [here](#)

Data submitted to EVA is brokered to our collaborating databases at [NCBI](#), [dbSNP](#) and [dbVar](#). It is therefore unnecessary to submit data to multiple resources. Please contact [eva-helpdesk@ebi.ac.uk](mailto:eva-helpdesk@ebi.ac.uk) if you would like any further information on this brokering process or collaboration.

If you have data in a format other than VCF, which cannot be converted to VCF, please contact [eva-helpdesk@ebi.ac.uk](mailto:eva-helpdesk@ebi.ac.uk). Additional submission formats may be supported over time as required by the scientific community.

### Key stages of EVA submissions

#### Contact

Contact [eva-helpdesk@ebi.ac.uk](mailto:eva-helpdesk@ebi.ac.uk) in order to provide details of your submission.

#### Receive

Download your [submission pack](#), which will include:

- Details for your submission uploads
- [Templates](#) to capture your associated metadata
- Key stages for your submission

#### Submit

Upload your data files to your private submission upload account or directly to the [eva-helpdesk@ebi.ac.uk](mailto:eva-helpdesk@ebi.ac.uk).

# Submission to EVA

- Data submitted to EVA is shared with NCBI:
  - dbSNP, dbVar
  - No need to submit data twice
- Response time of 48 hrs
  - Accession number suitable for publication
- EVA dynamic study loading pipeline
  - No need to wait to release date to see study on website or data in variant browser
- Focus on growth of PAG relevant data



# Conclusion



European Variation Archive  
[www.ebi.ac.uk/eva](http://www.ebi.ac.uk/eva)

- Variant file archive
- All types of variants, all species
- Provides direct views of raw VCF files
- Full API
- Completely free to use

# Acknowledgments

## EVA / DGVA

Justin Paschall

Ignacio Medina Castello

Gary Saunders

Cristina Yenyxe Gonzalez

Jag Kandasamy

Ilkka Lappalainen

## EGA

Jeff Almeida-King

Vasudev Kumanduri

Saif Ur-Rehman

Tom Smith



@ebivariation

## Ensembl Variation

Fiona Cunningham

Sarah Hunt

William McLaren

Anja Thormann

Laurent Gil

## ENA team

Rasko Leinonen

Rajesh Radhakrishnan

Daniel Vaughan

## Ensembl Genomes

Paul Kersey

Dan Bolser

Christoph Grabmuller

## Funding

