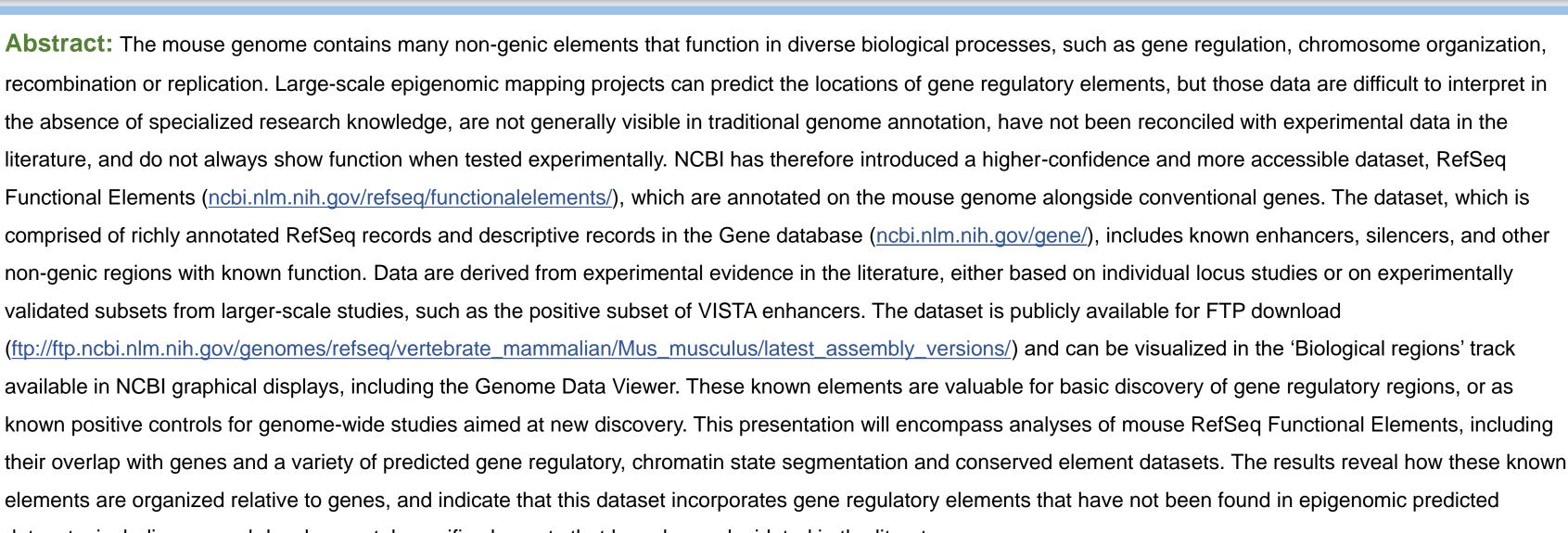


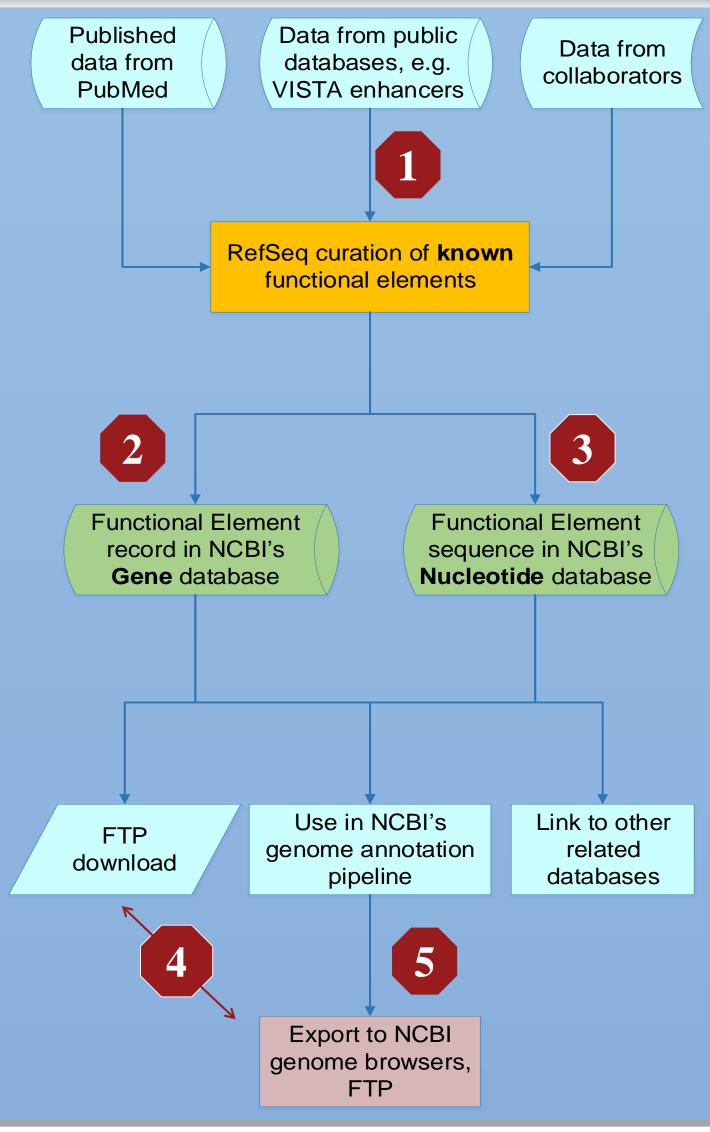
Poster #888C **Presentation on** April 28, 2020 12:02-12:04 PM

# Known versus Predicted: RefSeq Functional Elements as a **Reference Set of High-Confidence Non-Genic Elements in Mouse**

Catherine M. Farrell, Terence D. Murphy, and the RefSeq Curation and Development Team\* National Center for Biotechnology Information (NCBI), National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA



datasets, including several developmental-specific elements that have been elucidated in the literature.



### **RefSeq Functional Elements scope**

#### **Definition:**

Any non-genic genomic element that has functional significance based on experimental support

#### Types:

- Gene regulatory elements, e.g. enhancers, silencers, promoters, protein binding sites
- Known structural elements, e.g. boundary elements, chromatin conformationassociated regions
- Other elements of functional importance, e.g. well-defined recombination hotspots or replication origins



Follow NCBI on:

you the in

🔊 f 🗹

# **RefSeq records**

- Genomic accessions with NG\_ prefixes
- Data standard ontologies from:
  - International Nucleotide Sequence Database Collaboration (INSDC)
  - Sequence Ontology (SO)
  - Evidence & Conclusion Ontology (ECO)
- Links to publications in PubMed
- INSDC qualifiers with functional and descriptive metadata  $\succ$

### Example: Segment of NG\_055003.1, Hbb-ar, GeneID:15128

regulatory	388596 /regulatory_class="promoter" /experiment="EXISTENCE:reporter gene assay evidence [ECO:0000049][PMID: <u>12920040]</u> "			
	/note="-158 to +52 epsilon-y promoter fragment" /function="Hbb-y gene minimal promoter in transfected MEL			
	cells" /db_xref="GeneID: <u>15128</u> "			
regulatory	540596			
	/regulatory_class="transcriptional_cis_regulatory_region" /experiment="EXISTENCE:reporter gene assay evidence [ECO:0000049][PMID: <u>12920040]</u> "			
	/note="sequence between -100 and -158 that includes a			
	CACCC motif; based on difference between the activities of			

Annotation scope:

- Elements that have been experimentally validated
- Human and mouse elements

# Gene database records

#### ncbi.nlm.nih.gov/gene/

						Table of contents	
						Summary	
b-ar nemoglobir	n, activating	region [ <i>Mus musculus</i> (hou	ise mouse) j			Genomic context	
ne ID: 15128, updated on	12-Aug-2019					Genomic regions, transcripts, and products	
Summary					\$ ?	Bibliography	
,							
Official Symbol	Hbb-ar provided b	y <u>MGI</u>			Other	Variation	
		ating region provided by MGI				Interactions	
Primary source	MGI:MGI:102559	<u>l</u>			curated	General gene information	
	biological region				content	Homology	
Feature type(s)	protein_bind	a l hypersepsitive site enhancer enhan	eer blocking clome	at leave central region promotor	oomon	NCBI Reference Sequences (RefSeq)	
		e_l_hypersensitive_site, enhancer, enhan s_regulatory_region	cer_blocking_elemei	nt, locus_control_region, promoter,		Related sequences	
RefSeg status		S_regulatory_region					
Organism	Mus musculus					Additional links	
Lineage	Eukaryota; Metaz	zoa; Chordata; Craniata; Vertebrata; Eutel	eostomi; Mammalia;	Eutheria; Euarchontoglires; Glires; Roder	ntia; Myomorpha;		
		ae; Murinae; Mus; Mus				Genome Browsers	
Also known as		wanes which spans 24 kb unstream of th	a hamaslahin V hat	like embruenis chain (Libb v) sone, sons	inconto the El	Genome Data Viewer	
Summary	-	quence, which spans 24 kb upstream of the of the mouse beta-globin gene cluster on					
					-		
				ed on the chromosome in their order of de	evelopmental		
		genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb		ed on the chromosome in their order of de bb-bt in the 'single' haplotype (as represe		Related information	
	expression. The		o-bh3, Hbb-bs and H	bb-bt in the 'single' haplotype (as represe	nted in the	Related information BioProjects	
	expression. The GRCm38 primary 'diffuse' haplotype	genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb y reference assembly C57BL/6J strain alle e (as represented in the GRCm38 alternat	o-bh3, Hbb-bs and H le), or Hbb-y, Hbb-bh le BALB/c strain allel	bb-bt in the 'single' haplotype (as represe 10, Hbb-bh1, Hbb-bh2, Hbb-bh3, Hbb-b1 e). This regulatory region also includes th	nted in the and Hbb-b2 in the Hbb-y promoter		
	expression. The GRCm38 primary 'diffuse' haplotype and several DNa	genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb / reference assembly C57BL/6J strain alle e (as represented in the GRCm38 alternati se I hypersensitive sites (HSs; 5'HS1-5'HS	o-bh3, Hbb-bs and H le), or Hbb-y, Hbb-bh le BALB/c strain allel S6 in the entire region	bb-bt in the 'single' haplotype (as represent 10, Hbb-bh1, Hbb-bh2, Hbb-bh3, Hbb-b1 e). This regulatory region also includes th 1, 5'HS1-5'HS5 in the LCR), some of which	nted in the and Hbb-b2 in the le Hbb-y promoter ch possess mild	BioProjects Full text in PMC	
	expression. The g GRCm38 primary 'diffuse' haplotype and several DNa enhancer activity	genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb / reference assembly C57BL/6J strain alle e (as represented in the GRCm38 alternat se I hypersensitive sites (HSs; 5'HS1-5'HS in erythroid cells. The HSs bind to erythro	o-bh3, Hbb-bs and Hi lele), or Hbb-y, Hbb-bl le BALB/c strain allel S6 in the entire region oid transcription facto	bb-bt in the 'single' haplotype (as represent 0, Hbb-bh1, Hbb-bh2, Hbb-bh3, Hbb-b1 e). This regulatory region also includes th n, 5'HS1-5'HS5 in the LCR), some of which rs and are associated with regulatory con	nted in the and Hbb-b2 in the ie Hbb-y promoter ch possess mild nplexes involved in	BioProjects Full text in PMC Full text in PMC_nucleotide	
	expression. The g GRCm38 primary 'diffuse' haplotype and several DNa enhancer activity beta-globin gene	genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb / reference assembly C57BL/6J strain alle e (as represented in the GRCm38 alternati se I hypersensitive sites (HSs; 5'HS1-5'HS	b-bh3, Hbb-bs and Hi le), or Hbb-y, Hbb-bt te BALB/c strain allel S6 in the entire region bid transcription facto chromatin domain str	bb-bt in the 'single' haplotype (as represent 0, Hbb-bh1, Hbb-bh2, Hbb-bh3, Hbb-b1 e). This regulatory region also includes th n, 5'HS1-5'HS5 in the LCR), some of which rs and are associated with regulatory con ucture. This region also includes a weak (	nted in the and Hbb-b2 in the ie Hbb-y promoter ch possess mild nplexes involved in enhancer-blocking	BioProjects Full text in PMC	
	expression. The g GRCm38 primary 'diffuse' haplotype and several DNa enhancer activity beta-globin gene element at 5'HS5	genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb / reference assembly C57BL/6J strain alle e (as represented in the GRCm38 alternat se I hypersensitive sites (HSs; 5'HS1-5'HS in erythroid cells. The HSs bind to erythro expression and the organization of local c	b-bh3, Hbb-bs and Hi ele), or Hbb-y, Hbb-bt te BALB/c strain allel S6 in the entire region bid transcription facto chromatin domain str ays show that the LC	bb-bt in the 'single' haplotype (as represe 10, Hbb-bh1, Hbb-bh2, Hbb-bh3, Hbb-b1 e). This regulatory region also includes th n, 5'HS1-5'HS5 in the LCR), some of whic rs and are associated with regulatory con ucture. This region also includes a weak o R is an integral part of the beta-globin act	nted in the and Hbb-b2 in the le Hbb-y promoter ch possess mild nplexes involved in enhancer-blocking tive chromatin hub	BioProjects Full text in PMC Full text in PMC_nucleotide	
	expression. The g GRCm38 primary 'diffuse' haplotype and several DNa enhancer activity beta-globin gene element at 5'HS5 in erythroid cells, gene promoters.	genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb y reference assembly C57BL/6J strain alle e (as represented in the GRCm38 alternati se I hypersensitive sites (HSs; 5'HS1-5'HS in erythroid cells. The HSs bind to erythro expression and the organization of local c c. Chromosome conformation capture assa where it undergoes looping interactions w Mutations in the orthologous human LCR	b-bh3, Hbb-bs and Hi ele), or Hbb-y, Hbb-bf te BALB/c strain allel S6 in the entire region old transcription facto chromatin domain str ays show that the LC vith distal 5' and 3' His result in beta-thalass	bb-bt in the 'single' haplotype (as represent 0, Hbb-bh1, Hbb-bh2, Hbb-bh3, Hbb-b1 e). This regulatory region also includes the n, 5'HS1-5'HS5 in the LCR), some of which rs and are associated with regulatory con ucture. This region also includes a weak of R is an integral part of the beta-globin act Ss (HS-62.5, HS-60.7 and 3'HS1) and the semias and beta-hemoglobinopathies, and	nted in the and Hbb-b2 in the e Hbb-y promoter ch possess mild nplexes involved in enhancer-blocking tive chromatin hub e active beta-globin	BioProjects Full text in PMC Full text in PMC_nucleotide Gene neighbors	
	expression. The g GRCm38 primary 'diffuse' haplotype and several DNa enhancer activity beta-globin gene element at 5'HS5 in erythroid cells, gene promoters. region has been	genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb y reference assembly C57BL/6J strain alle e (as represented in the GRCm38 alternati se I hypersensitive sites (HSs; 5'HS1-5'HS in erythroid cells. The HSs bind to erythro expression and the organization of local c i. Chromosome conformation capture assa where it undergoes looping interactions w	b-bh3, Hbb-bs and Hi ele), or Hbb-y, Hbb-bf te BALB/c strain allel S6 in the entire region old transcription facto chromatin domain str ays show that the LC vith distal 5' and 3' His result in beta-thalass	bb-bt in the 'single' haplotype (as represent 0, Hbb-bh1, Hbb-bh2, Hbb-bh3, Hbb-b1 e). This regulatory region also includes the n, 5'HS1-5'HS5 in the LCR), some of which rs and are associated with regulatory con ucture. This region also includes a weak of R is an integral part of the beta-globin act Ss (HS-62.5, HS-60.7 and 3'HS1) and the semias and beta-hemoglobinopathies, and	nted in the and Hbb-b2 in the e Hbb-y promoter ch possess mild nplexes involved in enhancer-blocking tive chromatin hub e active beta-globin	BioProjects Full text in PMC Full text in PMC_nucleotide Gene neighbors Nucleotide PubMed	
Orthologs	expression. The g GRCm38 primary 'diffuse' haplotype and several DNa enhancer activity beta-globin gene element at 5'HS5 in erythroid cells, gene promoters. region has been	genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb y reference assembly C57BL/6J strain alle e (as represented in the GRCm38 alternati se I hypersensitive sites (HSs; 5'HS1-5'HS in erythroid cells. The HSs bind to erythro expression and the organization of local c c. Chromosome conformation capture assa where it undergoes looping interactions w Mutations in the orthologous human LCR	b-bh3, Hbb-bs and Hi ele), or Hbb-y, Hbb-bf te BALB/c strain allel S6 in the entire region old transcription facto chromatin domain str ays show that the LC vith distal 5' and 3' His result in beta-thalass	bb-bt in the 'single' haplotype (as represent 0, Hbb-bh1, Hbb-bh2, Hbb-bh3, Hbb-b1 e). This regulatory region also includes the n, 5'HS1-5'HS5 in the LCR), some of which rs and are associated with regulatory con ucture. This region also includes a weak of R is an integral part of the beta-globin act Ss (HS-62.5, HS-60.7 and 3'HS1) and the semias and beta-hemoglobinopathies, and	nted in the and Hbb-b2 in the e Hbb-y promoter ch possess mild nplexes involved in enhancer-blocking tive chromatin hub e active beta-globin	BioProjects Full text in PMC Full text in PMC_nucleotide Gene neighbors Nucleotide PubMed PubMed(nucleotide/PMC)	
	expression. The g GRCm38 primary 'diffuse' haplotype and several DNa enhancer activity beta-globin gene element at 5'HS5 in erythroid cells, gene promoters. region has been	genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb y reference assembly C57BL/6J strain alle e (as represented in the GRCm38 alternati se I hypersensitive sites (HSs; 5'HS1-5'HS in erythroid cells. The HSs bind to erythro expression and the organization of local c c. Chromosome conformation capture assa where it undergoes looping interactions w Mutations in the orthologous human LCR	b-bh3, Hbb-bs and Hi ele), or Hbb-y, Hbb-bf te BALB/c strain allel S6 in the entire region old transcription facto chromatin domain str ays show that the LC vith distal 5' and 3' His result in beta-thalass	bb-bt in the 'single' haplotype (as represent 0, Hbb-bh1, Hbb-bh2, Hbb-bh3, Hbb-b1 e). This regulatory region also includes the n, 5'HS1-5'HS5 in the LCR), some of which rs and are associated with regulatory con ucture. This region also includes a weak of R is an integral part of the beta-globin act Ss (HS-62.5, HS-60.7 and 3'HS1) and the semias and beta-hemoglobinopathies, and	nted in the and Hbb-b2 in the le Hbb-y promoter ch possess mild nplexes involved in enhancer-blocking tive chromatin hub e active beta-globin d thus this mouse	BioProjects Full text in PMC Full text in PMC_nucleotide Gene neighbors Nucleotide PubMed	
Orthologs Genomic context	expression. The g GRCm38 primary 'diffuse' haplotype and several DNa enhancer activity beta-globin gene element at 5'HS5 in erythroid cells, gene promoters. region has been	genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb y reference assembly C57BL/6J strain alle e (as represented in the GRCm38 alternati se I hypersensitive sites (HSs; 5'HS1-5'HS in erythroid cells. The HSs bind to erythro expression and the organization of local c c. Chromosome conformation capture assa where it undergoes looping interactions w Mutations in the orthologous human LCR	b-bh3, Hbb-bs and Hi ele), or Hbb-y, Hbb-bf te BALB/c strain allel S6 in the entire region old transcription facto chromatin domain str ays show that the LC vith distal 5' and 3' His result in beta-thalass	bb-bt in the 'single' haplotype (as represent 0, Hbb-bh1, Hbb-bh2, Hbb-bh3, Hbb-b1 e). This regulatory region also includes the n, 5'HS1-5'HS5 in the LCR), some of which rs and are associated with regulatory con ucture. This region also includes a weak of R is an integral part of the beta-globin act Ss (HS-62.5, HS-60.7 and 3'HS1) and the semias and beta-hemoglobinopathies, and	nted in the and Hbb-b2 in the e Hbb-y promoter ch possess mild nplexes involved in enhancer-blocking tive chromatin hub e active beta-globin	BioProjects Full text in PMC Full text in PMC_nucleotide Gene neighbors Nucleotide PubMed PubMed(nucleotide/PMC)	
	expression. The g GRCm38 primary 'diffuse' haplotype and several DNa enhancer activity beta-globin gene element at 5'HS5 in erythroid cells, gene promoters. region has been	genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb y reference assembly C57BL/6J strain alle e (as represented in the GRCm38 alternati se I hypersensitive sites (HSs; 5'HS1-5'HS in erythroid cells. The HSs bind to erythro expression and the organization of local c c. Chromosome conformation capture assa where it undergoes looping interactions w Mutations in the orthologous human LCR	b-bh3, Hbb-bs and Hi ele), or Hbb-y, Hbb-bf te BALB/c strain allel S6 in the entire region old transcription facto chromatin domain str ays show that the LC vith distal 5' and 3' His result in beta-thalass	bb-bt in the 'single' haplotype (as represe 10, Hbb-bh1, Hbb-bh2, Hbb-bh3, Hbb-b1 e). This regulatory region also includes th n, 5'HS1-5'HS5 in the LCR), some of which rs and are associated with regulatory con- ucture. This region also includes a weak of R is an integral part of the beta-globin act SS (HS-62.5, HS-60.7 and 3'HS1) and the semias and beta-hemoglobinopathies, and [provided by RefSeq, Aug 2017]	nted in the and Hbb-b2 in the le Hbb-y promoter ch possess mild nplexes involved in enhancer-blocking tive chromatin hub e active beta-globin d thus this mouse	BioProjects Full text in PMC Full text in PMC_nucleotide Gene neighbors Nucleotide PubMed PubMed(nucleotide/PMC) Taxonomy	
Genomic context	expression. The g GRCm38 primary 'diffuse' haplotype and several DNa enhancer activity beta-globin gene element at 5'HS5 in erythroid cells, gene promoters. region has been	genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb y reference assembly C57BL/6J strain alle e (as represented in the GRCm38 alternati se I hypersensitive sites (HSs; 5'HS1-5'HS in erythroid cells. The HSs bind to erythro expression and the organization of local c c. Chromosome conformation capture assa where it undergoes looping interactions w Mutations in the orthologous human LCR	b-bh3, Hbb-bs and Hi ele), or Hbb-y, Hbb-bf te BALB/c strain allel S6 in the entire region old transcription facto chromatin domain str ays show that the LC vith distal 5' and 3' His result in beta-thalass	bb-bt in the 'single' haplotype (as represe 10, Hbb-bh1, Hbb-bh2, Hbb-bh3, Hbb-b1 e). This regulatory region also includes th n, 5'HS1-5'HS5 in the LCR), some of which rs and are associated with regulatory con- ucture. This region also includes a weak of R is an integral part of the beta-globin act SS (HS-62.5, HS-60.7 and 3'HS1) and the semias and beta-hemoglobinopathies, and [provided by RefSeq, Aug 2017]	nted in the and Hbb-b2 in the le Hbb-y promoter ch possess mild mplexes involved in enhancer-blocking tive chromatin hub e active beta-globin d thus this mouse	BioProjects Full text in PMC Full text in PMC_nucleotide Gene neighbors Nucleotide PubMed PubMed(nucleotide/PMC) Taxonomy	
Genomic context	expression. The g GRCm38 primary 'diffuse' haplotype and several DNa enhancer activity beta-globin gene element at 5'HS5 in erythroid cells, gene promoters. region has been	genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb y reference assembly C57BL/6J strain alle e (as represented in the GRCm38 alternati se I hypersensitive sites (HSs; 5'HS1-5'HS in erythroid cells. The HSs bind to erythro expression and the organization of local c c. Chromosome conformation capture assa where it undergoes looping interactions w Mutations in the orthologous human LCR	b-bh3, Hbb-bs and Hi ele), or Hbb-y, Hbb-bf te BALB/c strain allel S6 in the entire region old transcription facto chromatin domain str ays show that the LC vith distal 5' and 3' His result in beta-thalass	bb-bt in the 'single' haplotype (as represe 10, Hbb-bh1, Hbb-bh2, Hbb-bh3, Hbb-b1 e). This regulatory region also includes th n, 5'HS1-5'HS5 in the LCR), some of which rs and are associated with regulatory con- ucture. This region also includes a weak of R is an integral part of the beta-globin act SS (HS-62.5, HS-60.7 and 3'HS1) and the semias and beta-hemoglobinopathies, and [provided by RefSeq, Aug 2017]	nted in the and Hbb-b2 in the le Hbb-y promoter ch possess mild mplexes involved in enhancer-blocking tive chromatin hub e active beta-globin d thus this mouse	BioProjects Full text in PMC Full text in PMC_nucleotide Gene neighbors Nucleotide PubMed PubMed(nucleotide/PMC) Taxonomy Links to other resources	
Genomic context	expression. The g GRCm38 primary 'diffuse' haplotype and several DNa enhancer activity beta-globin gene element at 5'HS5 in erythroid cells, gene promoters. region has been human all	genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb v reference assembly C57BL/6J strain alle e (as represented in the GRCm38 alternati- se I hypersensitive sites (HSs; 5'HS1-5'HS in erythroid cells. The HSs bind to erythro- expression and the organization of local of c. Chromosome conformation capture assa- where it undergoes looping interactions w Mutations in the orthologous human LCR intensively studied as a model for beta-glo	-bh3, Hbb-bs and H le), or Hbb-y, Hbb-bf le BALB/c strain allel 56 in the entire region bid transcription factory chromatin domain str ays show that the LC vith distal 5' and 3' H result in beta-thalass bbin gene regulation.	bb-bt in the 'single' haplotype (as represe 10, Hbb-bh1, Hbb-bh2, Hbb-bh3, Hbb-b1 e). This regulatory region also includes th h, 5'HS1-5'HS5 in the LCR), some of which rs and are associated with regulatory con- ucture. This region also includes a weak of R is an integral part of the beta-globin act Ss (HS-62.5, HS-60.7 and 3'HS1) and the semias and beta-hemoglobinopathies, and [provided by RefSeq, Aug 2017] See Hbb-ar in	nted in the and Hbb-b2 in the the Hbb-y promoter th possess mild mplexes involved in enhancer-blocking tive chromatin hub a active beta-globin d thus this mouse	BioProjects Full text in PMC Full text in PMC_nucleotide Gene neighbors Nucleotide PubMed PubMed(nucleotide/PMC) Taxonomy Links to other resources MGI	
Genomic context ocation: 7 E2; 7 Annotation release	expression. The g GRCm38 primary 'diffuse' haplotyp and several DNa enhancer activity beta-globin gene element at 5'HS5 in erythroid cells, gene promoters. region has been human all	genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb y reference assembly C57BL/6J strain alle e (as represented in the GRCm38 alternati se I hypersensitive sites (HSs; 5'HS1-5'HS in erythroid cells. The HSs bind to erythro expression and the organization of local of c. Chromosome conformation capture assa where it undergoes looping interactions w Mutations in the orthologous human LCR intensively studied as a model for beta-glo Assembly	-bh3, Hbb-bs and Hi de), or Hbb-y, Hbb-bf te BALB/c strain allel 56 in the entire region oid transcription facto chromatin domain str ays show that the LC vith distal 5' and 3' Hi result in beta-thalass obin gene regulation.	bb-bt in the 'single' haplotype (as represended), Hbb-bh1, Hbb-bh2, Hbb-bh3, Hbb-bh1, e). This regulatory region also includes the first and are associated with regulatory conducture. This region also includes a weak of R is an integral part of the beta-globin actions and beta-hemoglobinopathies, and [provided by RefSeq, Aug 2017] See Hbb-ar in Location	nted in the and Hbb-b2 in the the Hbb-y promoter th possess mild mplexes involved in enhancer-blocking tive chromatin hub a active beta-globin d thus this mouse	BioProjects Full text in PMC Full text in PMC_nucleotide Gene neighbors Nucleotide PubMed PubMed(nucleotide/PMC) Taxonomy Links to other resources MGI General information	
Genomic context ocation: 7 E2; 7 Annotation release	expression. The g GRCm38 primary 'diffuse' haplotype and several DNa enhancer activity beta-globin gene element at 5'HS5 in erythroid cells, gene promoters. region has been human all Status current	genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb y reference assembly C57BL/6J strain alle e (as represented in the GRCm38 alternation se I hypersensitive sites (HSs; 5'HS1-5'HS in erythroid cells. The HSs bind to erythro expression and the organization of local of b. Chromosome conformation capture assa where it undergoes looping interactions with Mutations in the orthologous human LCR intensively studied as a model for beta-glo Assembly GRCm38.p6 (GCF_000001635.26) Chromosome 7	-bh3, Hbb-bs and Hi le), or Hbb-y, Hbb-bf le BALB/c strain allel 56 in the entire region old transcription facto chromatin domain str ays show that the LC vith distal 5' and 3' Hi result in beta-thalass obin gene regulation.	bb-bt in the 'single' haplotype (as represe 10, Hbb-bh1, Hbb-bh2, Hbb-bh3, Hbb-b1 e). This regulatory region also includes th n, 5'HS1-5'HS5 in the LCR), some of which rs and are associated with regulatory con- ucture. This region also includes a weak of R is an integral part of the beta-globin act SS (HS-62.5, HS-60.7 and 3'HS1) and the semias and beta-hemoglobinopathies, and [provided by RefSeq, Aug 2017] See Hbb-ar in Location NC_000073.6 (10385276610387754	nted in the and Hbb-b2 in the the Hbb-y promoter th possess mild mplexes involved in enhancer-blocking tive chromatin hub a active beta-globin d thus this mouse	BioProjects Full text in PMC Full text in PMC_nucleotide Gene neighbors Nucleotide PubMed PubMed(nucleotide/PMC) Taxonomy Links to other resources MGI General information About Gene	
Genomic context ocation: 7 E2; 7 Annotation release	expression. The g GRCm38 primary 'diffuse' haplotype and several DNa enhancer activity beta-globin gene element at 5'HS5 in erythroid cells, gene promoters. region has been human all Status current	genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb y reference assembly C57BL/6J strain alle e (as represented in the GRCm38 alternation se I hypersensitive sites (HSs; 5'HS1-5'HS in erythroid cells. The HSs bind to erythro expression and the organization of local co- c. Chromosome conformation capture assa where it undergoes looping interactions w Mutations in the orthologous human LCR intensively studied as a model for beta-glo Assembly GRCm38.p6 (GCF_000001635.26) Chromosome 7	-bh3, Hbb-bs and Hi le), or Hbb-y, Hbb-bf le BALB/c strain allel 56 in the entire region old transcription facto chromatin domain str ays show that the LC vith distal 5' and 3' Hi result in beta-thalass obin gene regulation. Chr 7 - NC_000073.6	bb-bt in the 'single' haplotype (as represe 10, Hbb-bh1, Hbb-bh2, Hbb-bh3, Hbb-b1 e). This regulatory region also includes th n, 5'HS1-5'HS5 in the LCR), some of whic res and are associated with regulatory con ucture. This region also includes a weak ( R is an integral part of the beta-globin act Ss (HS-62.5, HS-60.7 and 3'HS1) and the semias and beta-hemoglobinopathies, and [provided by RefSeq, Aug 2017] See Hbb-ar in Location NC_000073.6 (10385276610387754	nted in the and Hbb-b2 in the the Hbb-y promoter th possess mild mplexes involved in enhancer-blocking tive chromatin hub a active beta-globin d thus this mouse	BioProjects Full text in PMC Full text in PMC_nucleotide Gene neighbors Nucleotide PubMed PubMed(nucleotide/PMC) Taxonomy Links to other resources MGI General information About Gene FAQ	
Genomic context ocation: 7 E2; 7 Annotation release	expression. The g GRCm38 primary 'diffuse' haplotype and several DNa enhancer activity beta-globin gene element at 5'HS5 in erythroid cells, gene promoters. region has been human all Status current	genes are: Hbb-y, Hbb-bh1, Hbb-bh2, Hbb y reference assembly C57BL/6J strain alle e (as represented in the GRCm38 alternation se I hypersensitive sites (HSs; 5'HS1-5'HS in erythroid cells. The HSs bind to erythro expression and the organization of local of b. Chromosome conformation capture assa where it undergoes looping interactions with Mutations in the orthologous human LCR intensively studied as a model for beta-glo Assembly GRCm38.p6 (GCF_000001635.26) Chromosome 7	-bh3, Hbb-bs and Hi le), or Hbb-y, Hbb-bf le BALB/c strain allel 56 in the entire region old transcription facto chromatin domain str ays show that the LC vith distal 5' and 3' Hi result in beta-thalass obin gene regulation. Chr 7 - NC_000073.6	bb-bt in the 'single' haplotype (as represe 10, Hbb-bh1, Hbb-bh2, Hbb-bh3, Hbb-b1 e). This regulatory region also includes th n, 5'HS1-5'HS5 in the LCR), some of which rs and are associated with regulatory con- ucture. This region also includes a weak of R is an integral part of the beta-globin act SS (HS-62.5, HS-60.7 and 3'HS1) and the semias and beta-hemoglobinopathies, and [provided by RefSeq, Aug 2017] See Hbb-ar in Location NC_000073.6 (10385276610387754	nted in the and Hbb-b2 in the the Hbb-y promoter th possess mild mplexes involved in enhancer-blocking tive chromatin hub a active beta-globin d thus this mouse	BioProjects Full text in PMC Full text in PMC_nucleotide Gene neighbors Nucleotide PubMed PubMed(nucleotide/PMC) Taxonomy Links to other resources MGI General information About Gene	

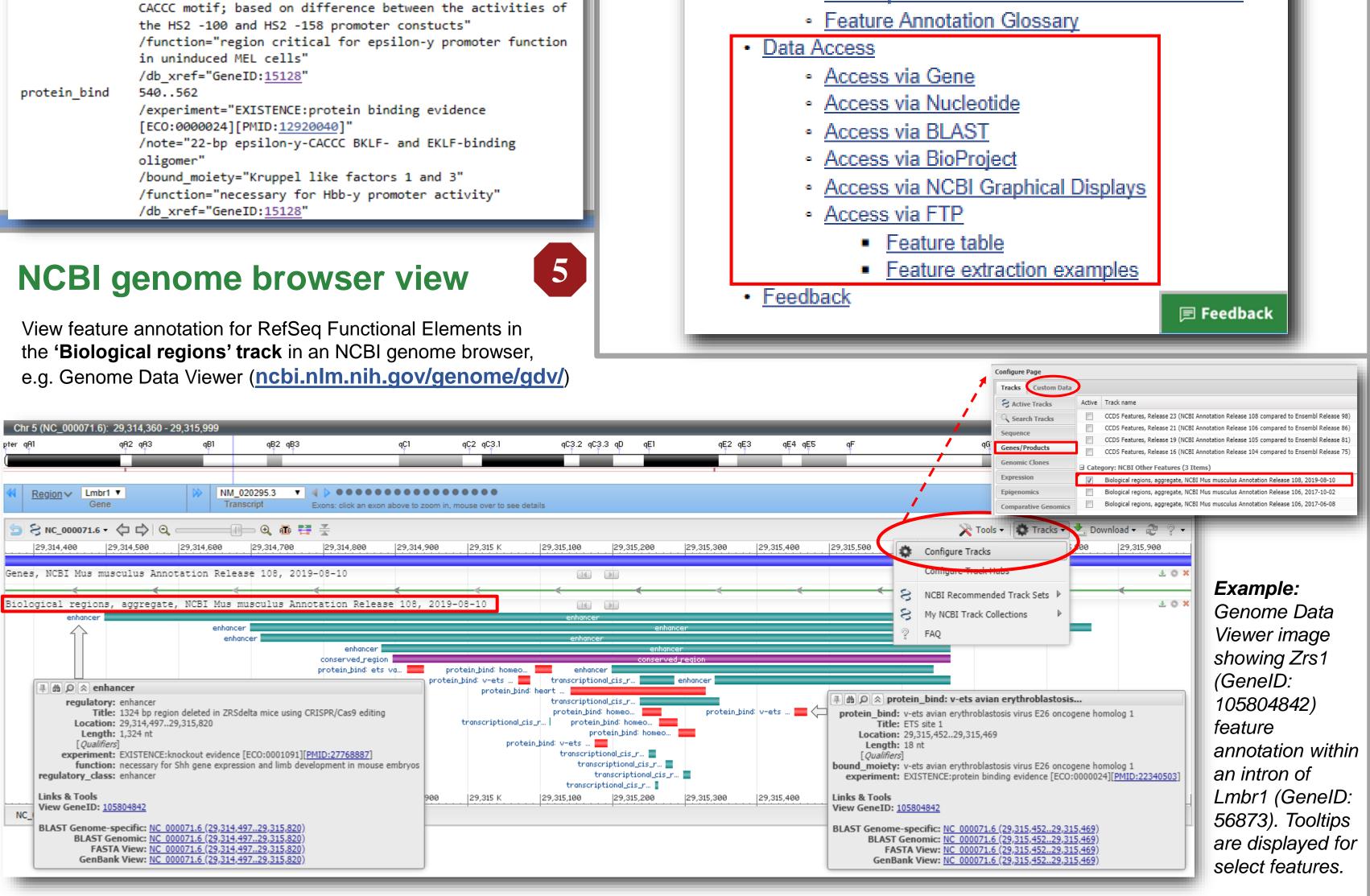
### Data access

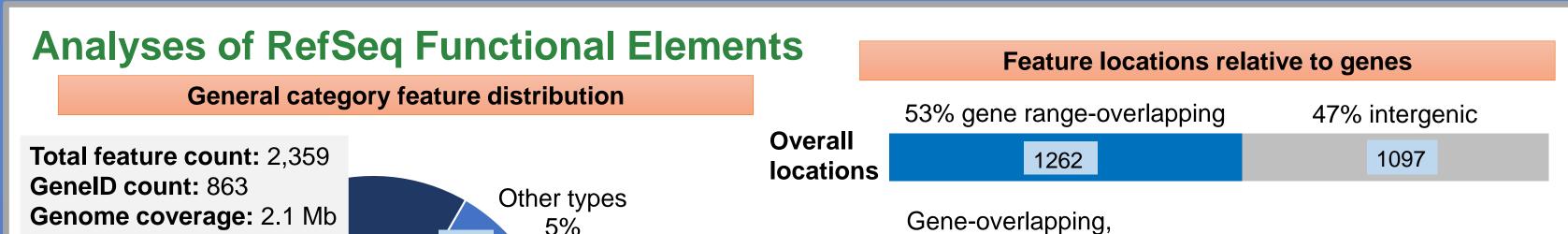
3

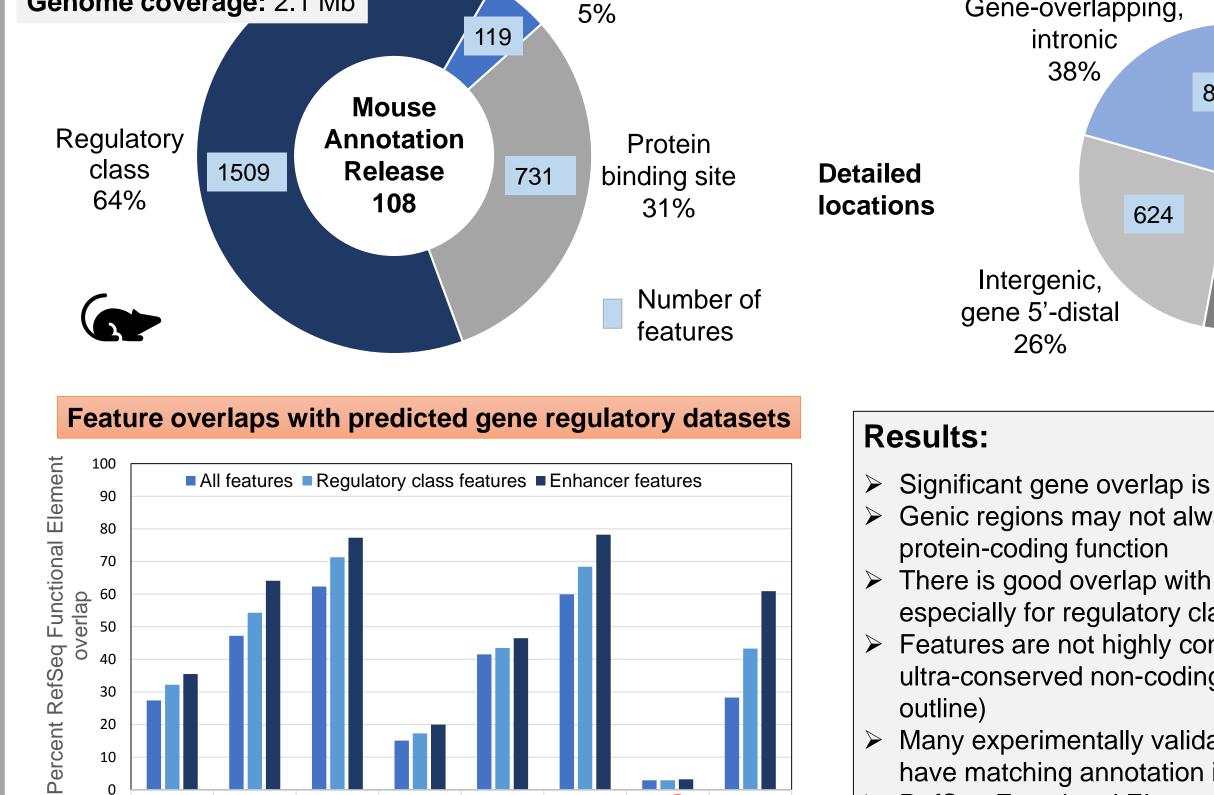
Find multiple ways to access the data on our website: ncbi.nlm.nih.gov/refseq/functionalelements/

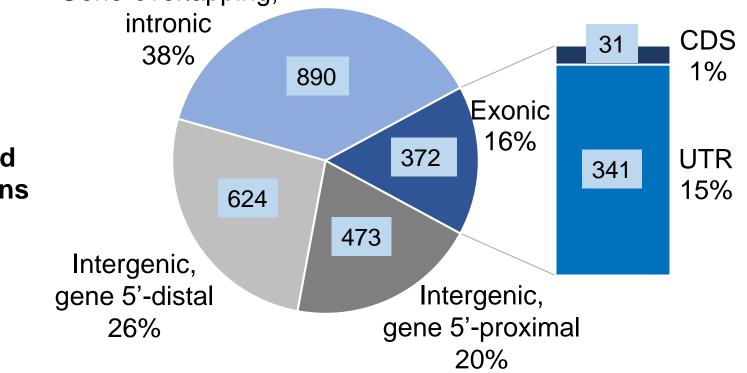
### **NCBI RefSeq Functional Elements**

- Overview
- RefSeg Functional Element Records
  - RefSeg Functional Element Feature Annotation









- Significant gene overlap is observed
- Genic regions may not always have exclusive transcript or
- $\succ$  There is good overlap with predicted gene regulatory datasets, especially for regulatory class and enhancer features
- Features are not highly conserved based on low overlap with ultra-conserved non-coding elements from UCNEbase (red outline)
- Many experimentally validated gene regulatory elements do not  $\geq$ have matching annotation in epigenomic predicted datasets
- RefSeq Functional Elements can be used for:
  - Discovery of non-coding genome function
  - As positive controls for large-scale epigenomic studies

### **Gene association caution:**

dbsUPER ODE 2-1.20 Ensemble 21 FANTON'S C

50

40

30

20

10

Gene regulatory elements don't necessarily regulate the genes they overlap or are closest to!

1CNED25e

- *Example:* the *Zrs1* limb enhancer in an intron of the *Lmbr1* gene (shown in browser image above) regulates *Shh* located 1 Mb away!
- ✤ Also check out RefSeq Functional Elements for human, and RefSeq gene/transcript/protein annotation!

VISTA

\* Do you have comments, suggestions or experimentally validated data to include in the dataset? Find me on TAGC2020 Slack, email me at farrelca@ncbi.nlm.nih.gov, or use the green Feedback button on our webpage, ncbi.nlm.nih.gov/refseq/functionalelements/

#### **J.S.** National Library of Medicine National Center for Biotechnology Information

8600 Rockville Pike, Bethesda MD, 20894 USA

This work was supported by the Intramural Research Program of the NIH, National Library of Medicine

\*RefSeq Curators: C.M. Farrell, T. Goldfarb, S.H. Rangwala RefSeq Developers: T.D. Murphy, A. Astashyn, O. Ermolaeva, V. Kodali, C. Wallin

This work was done as part of the Authors' official duties as NIH employees and is a Work of the United States Government. Therefore, copyright may not be established in the United States. 17 U.S.C. § 105