

# Transcriptomics Practical Setup

HackBio

---

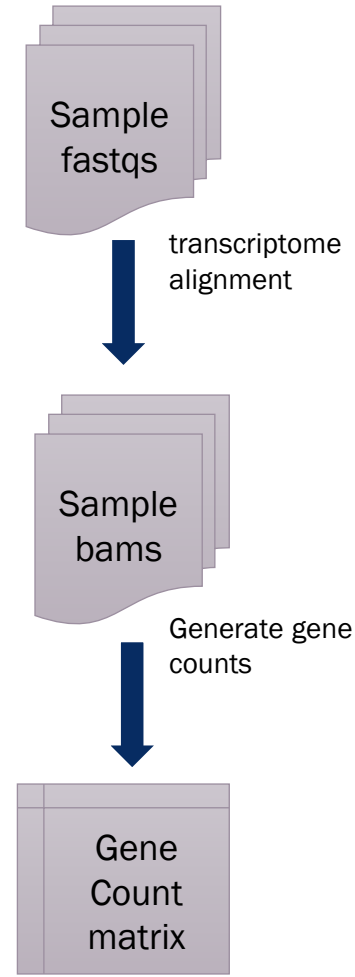
Melyssa Minto

West Lab, Duke Neurobiology

Computational Biology and Bioinformatics

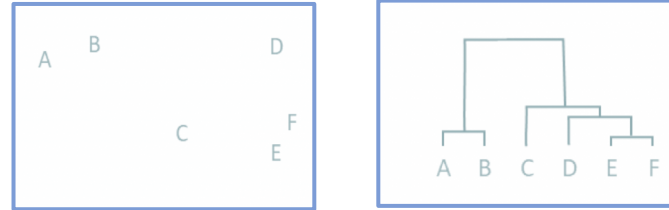
# Transcriptomics pipeline/workflow

## Preprocessing

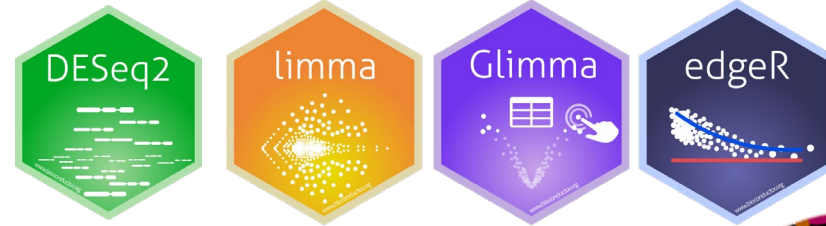
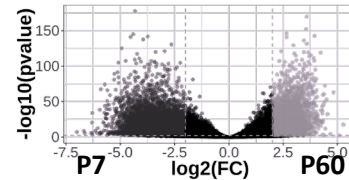


## Analyses

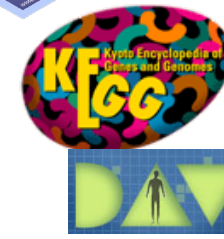
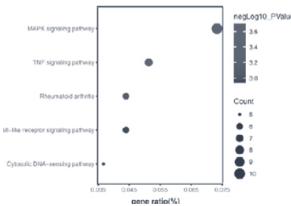
### Clustering



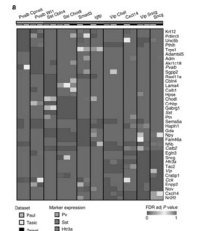
### Differential Expression



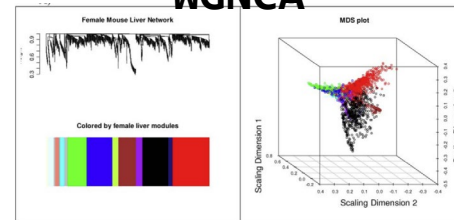
### Functional Enrichment



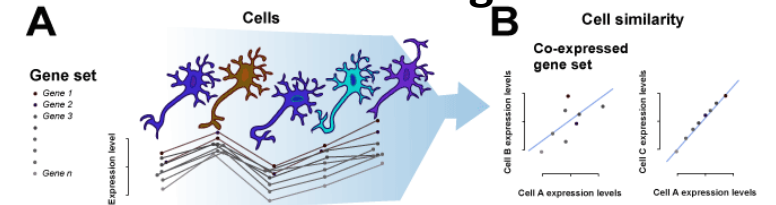
### Coregulated Gene Expression



### WGNCA

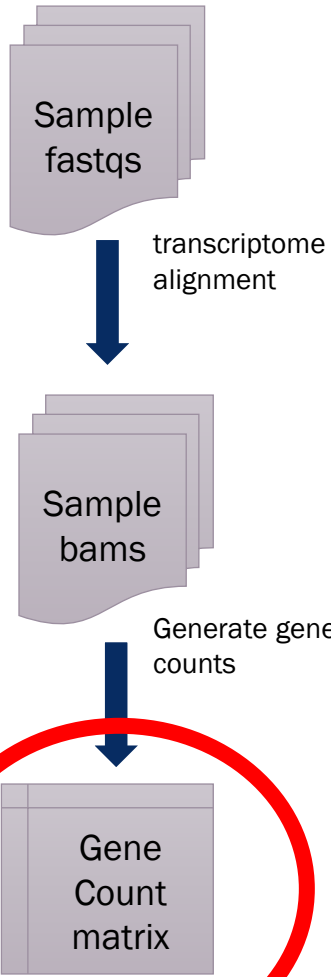


### MetaNeighbor



# Transcriptomics pipeline/workflow

## Preprocessing

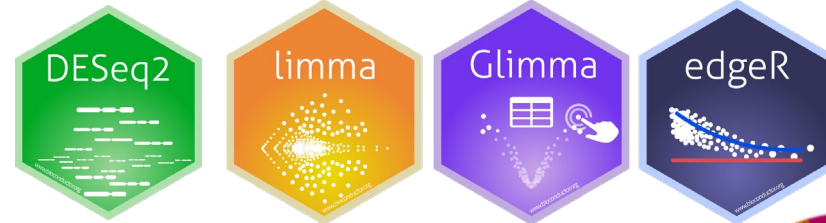
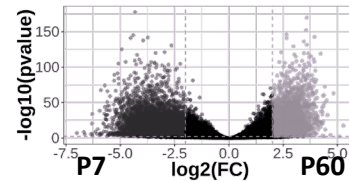


## Analyses

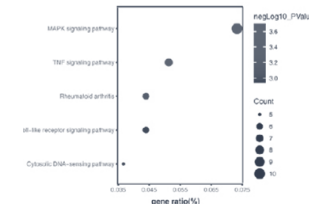
### Clustering



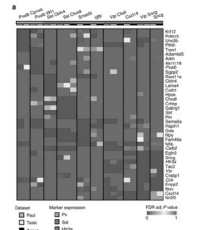
### Differential Expression



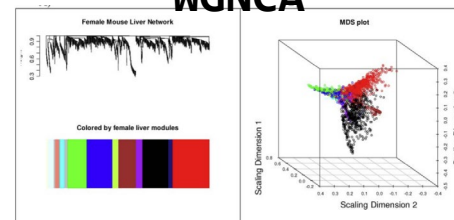
### Functional Enrichment



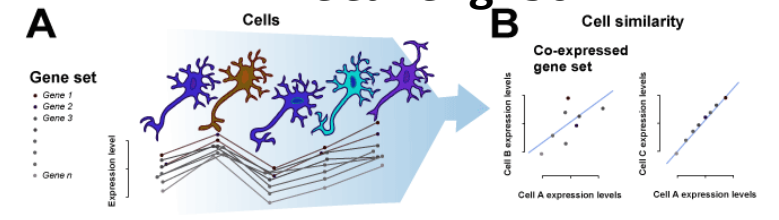
### Coregulated Gene Expression



### WGNCA



### MetaNeighbor








ARTICLE

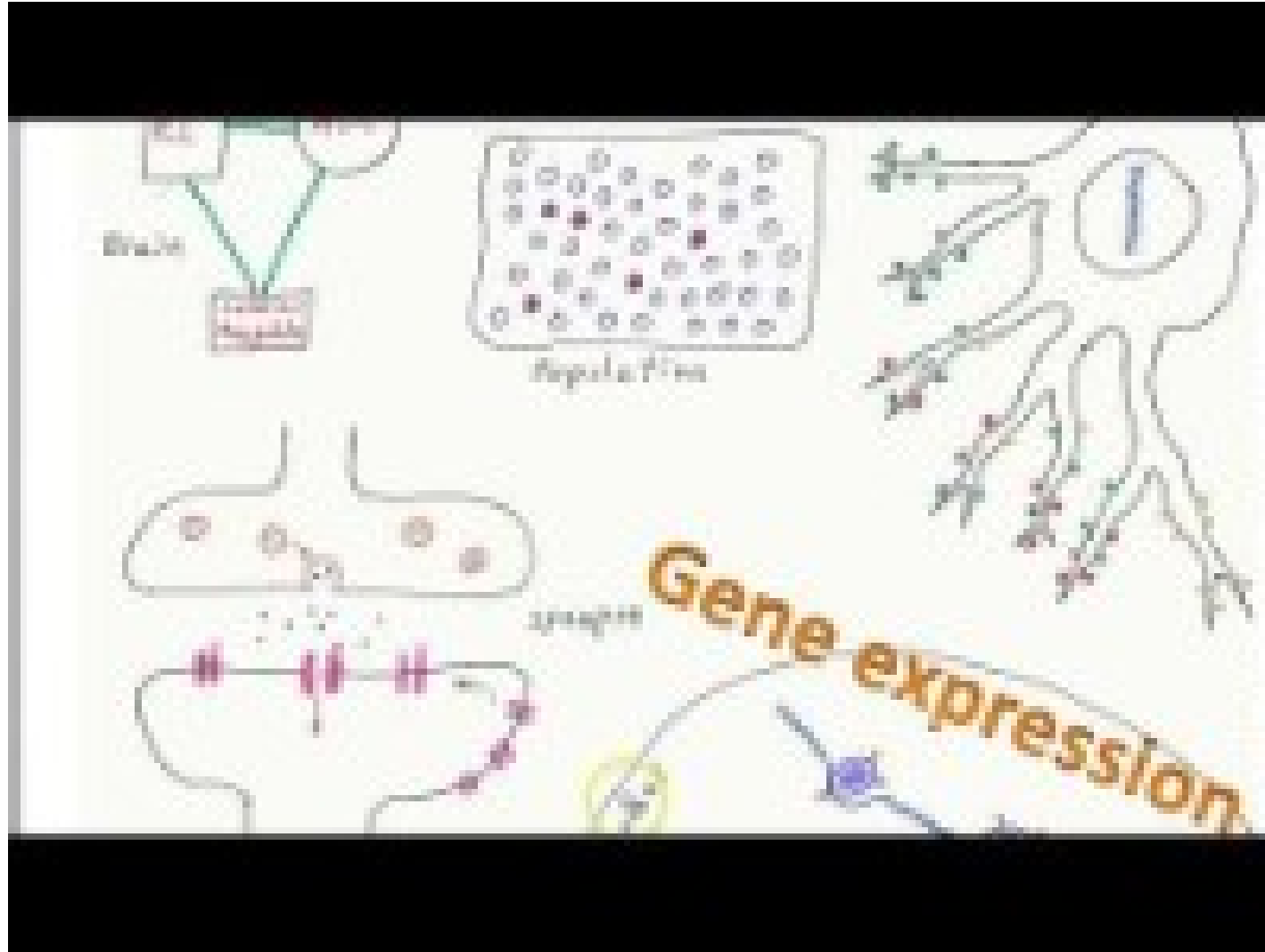
<https://doi.org/10.1038/s41467-019-09960-x>

OPEN

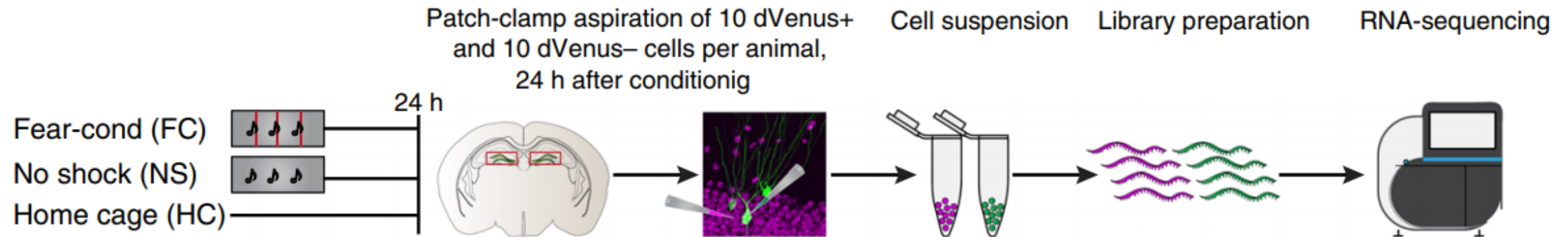
# Engram-specific transcriptome profiling of contextual memory consolidation

Priyanka Rao-Ruiz <sup>1,2</sup>, Jonathan J. Couey<sup>1</sup>, Ivo M. Marcelo<sup>1,3</sup>, Christian G. Bouwkamp<sup>1</sup>, Denise E. Slump<sup>1</sup>, Mariana R. Matos<sup>2</sup>, Rolinka J. van der Loo<sup>2</sup>, Gabriela J. Martins<sup>3,4</sup>, Mirjam van den Hout <sup>5</sup>, Wilfred F. van Ijcken <sup>5</sup>, Rui M. Costa<sup>3,4</sup>, Michel C. van den Oever <sup>2</sup> & Steven A. Kushner <sup>1</sup>

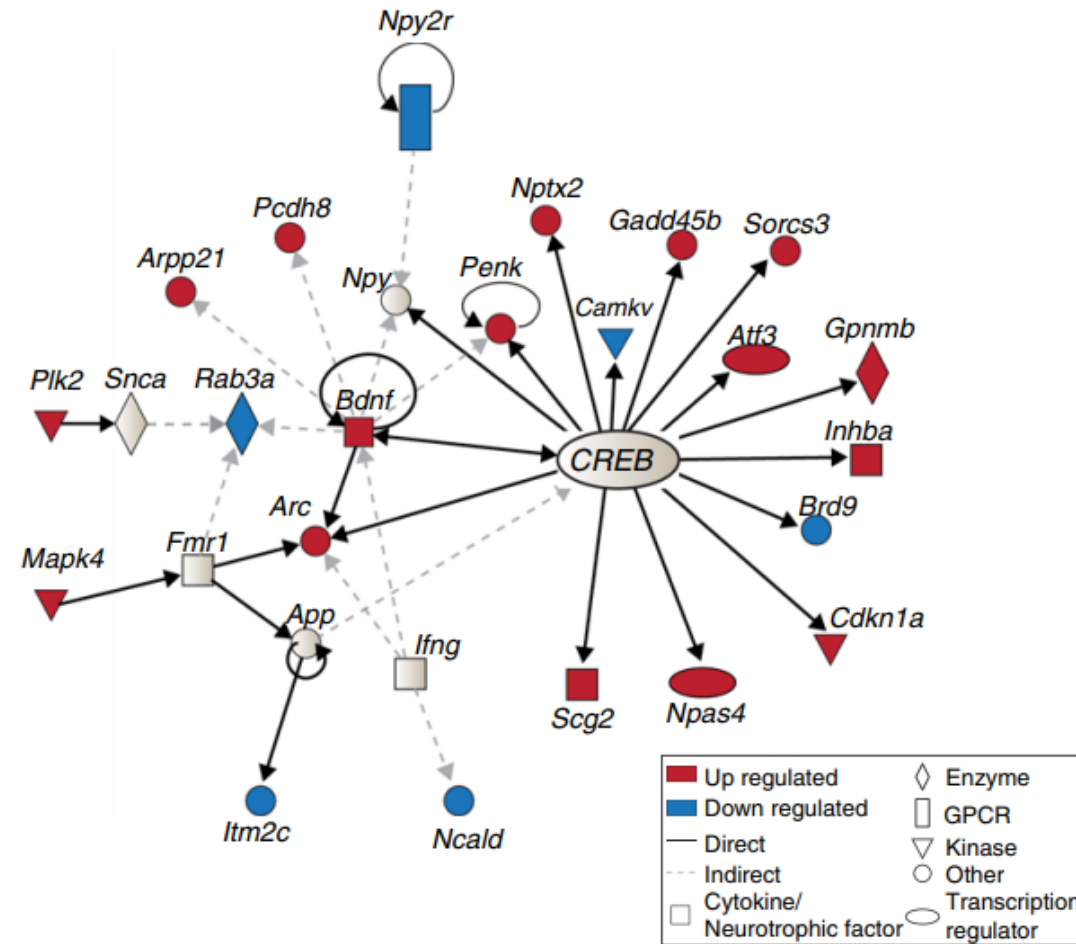
# What is an engram??





# Experimental design to isolate engram cells that participate in fear learning



# Novel target genes for CREB found that modulates fear memory learning



# Accessing the data through NCBI GEO

  
Gene Expression Omnibus

COVID-19 is an emerging, rapidly evolving situation.  
Get the latest public health information from CDC: <https://www.coronavirus.gov>.  
Get the latest research from NIH: <https://www.nih.gov/coronavirus>.  
Find NCBI SARS-CoV-2 literature, sequence, and clinical content: <https://www.ncbi.nlm.nih.gov/sars-cov-2/>.

[HOME](#) [SEARCH](#) [SITE MAP](#) [GEO Publications](#) [FAQ](#) [MIAME](#) [Email GEO](#)

[NCBI > GEO > Accession Display](#) [Not logged in | Login](#)

**GEO help:** Mouse over screen elements for information.

Scope:  Format:  Amount:  GEO accession:

**Series GSE129024** [Query DataSets for GSE129024](#)

Status Public on Mar 30, 2019

Title Engram-specific transcriptome profiling of contextual memory consolidation

Organism [Mus musculus](#)

Experiment type Expression profiling by high throughput sequencing

Summary Sparse populations of neurons in the dentate gyrus (DG) of the hippocampus are causally implicated in the encoding of contextual fear memories. However, engram-specific molecular mechanisms underlying memory consolidation remain largely unknown. Here we perform unbiased RNA sequencing of DG engram neurons 24h after contextual fear conditioning to identify transcriptome changes specific to memory consolidation. DG engram neurons exhibit a highly distinct pattern of gene expression, in which CREB-dependent transcription features prominently (P=6.2x10<sup>-13</sup>), including Atf3 (P=2.4x10<sup>-41</sup>), Penk (P=1.3x10<sup>-15</sup>), and Kcnq3 (P=3.1x10<sup>-12</sup>). Moreover, we validate the functional relevance of the RNAseq findings by establishing the causal requirement of intact CREB function specifically within the DG engram during memory consolidation, and identify a novel group of CREB target genes involved in the encoding of long-term memory.

Overall design Biological replicates: Fear conditioned: n=14, No shock controls: n=4, Home cage controls: n=3. The contents 10 dVenus+ and 10 dVenus- cells were aspirated from each animal (biological replicate)

Contributor(s) [Rao-Ruiz P, Couey JJ, Marcelo IM, Bouwkamp CG, Slump DE, Matos MR, van der Loo RJ, Martins GJ, van den Hout M, van IJcken WF, Costa RM, van den Oever MC, Kushner SA](#)

Citation(s) Rao-Ruiz P, Couey JJ, Marcelo IM, Bouwkamp CG et al. Engram-specific transcriptome profiling of contextual memory consolidation. *Nat Commun* 2019 May 20;10(1):2232. PMID: [31110186](#)

Submission date Mar 29, 2019

Last update date May 23, 2019

Contact name Steven Kushner

E-mail(s) [s.kushner@erasmusmc.nl](mailto:s.kushner@erasmusmc.nl)

Organization name Erasmus MC

[der Loo RJ, Martins GJ, van den Hout M, van IJcken WF, Costa RM, van den Oever MC, Kushner SA](#)

Citation(s) Rao-Ruiz P, Couey JJ, Marcelo IM, Bouwkamp CG et al. Engram-specific transcriptome profiling of contextual memory consolidation. *Nat Commun* 2019 May 20;10(1):2232. PMID: [31110186](#)

Submission date Mar 29, 2019

Last update date May 23, 2019

Contact name Steven Kushner

E-mail(s) [s.kushner@erasmusmc.nl](mailto:s.kushner@erasmusmc.nl)

Organization name Erasmus MC

Department Psychiatry

Street address Dr. Molewaterplein 40

City Rotterdam

State/province Zuid Holland

ZIP/Postal code 3015 GD

Country Netherlands

Platforms (1) [GPL17021](#) Illumina HiSeq 2500 (Mus musculus)

Samples (38) [GSM3690851](#) Fear Conditioned, dVenus+, MouseNr 01 [G01]  
[GSM3690852](#) Fear Conditioned, dVenus+, MouseNr 03 [G03]  
[GSM3690853](#) Fear Conditioned, dVenus-, MouseNr 01 [NG01]

**Relations**

BioProject [PRJNA529794](#)

SRA [SRP189843](#)

Download family	Format
<a href="#">SOFT formatted family file(s)</a>	SOFT <a href="#">?</a>
<a href="#">MINiML formatted family file(s)</a>	MINiML <a href="#">?</a>
<a href="#">Series Matrix File(s)</a>	TXT <a href="#">?</a>

Supplementary file	Size	Download	File type/resource
<a href="#">GSE129024_RAW.tar</a>	3.2 Mb	<a href="#">(http)</a> <a href="#">(custom)</a>	TAR (of TXT)
<a href="#">GSE129024_counts_per_gene_sample.txt.gz</a>	647.9 Kb	<a href="#">(ftp)</a> <a href="#">(http)</a>	TXT

[SRA Run Selector](#) [?](#)



Raw data are available in SRA

Processed data provided as supplementary file

Processed data are available on Series record



# Accessing the data through NCBI GEO

  
Gene Expression Omnibus

COVID-19 is an emerging, rapidly evolving situation.  
Get the latest public health information from CDC: <https://www.coronavirus.gov>.  
Get the latest research from NIH: <https://www.nih.gov/coronavirus>.  
Find NCBI SARS-CoV-2 literature, sequence, and clinical content: <https://www.ncbi.nlm.nih.gov/sars-cov-2/>.

[HOME](#) [SEARCH](#) [SITE MAP](#) [GEO Publications](#) [FAQ](#) [MIAME](#) [Email GEO](#)

[NCBI > GEO > Accession Display](#) [Not logged in | Login](#)

GEO help: Mouse over screen elements for information.

Scope:  Format:  Amount:  GEO accession:

**Series GSE129024** [Query DataSets for GSE129024](#)

Status Public on Mar 30, 2019  
Title Engram-specific transcriptome profiling of contextual memory consolidation  
Organism [Mus musculus](#)  
Experiment type Expression profiling by high throughput sequencing  
Summary Sparse populations of neurons in the dentate gyrus (DG) of the hippocampus are causally implicated in the encoding of contextual fear memories. However, engram-specific molecular mechanisms underlying memory consolidation remain largely unknown. Here we perform unbiased RNA sequencing of DG engram neurons 24h after contextual fear conditioning to identify transcriptome changes specific to memory consolidation. DG engram neurons exhibit a highly distinct pattern of gene expression, in which CREB-dependent transcription features prominently (P=6.2x10<sup>-13</sup>), including Atf3 (P=2.4x10<sup>-41</sup>), Penk (P=1.3x10<sup>-15</sup>), and Kcnq3 (P=3.1x10<sup>-12</sup>). Moreover, we validate the functional relevance of the RNAseq findings by establishing the causal requirement of intact CREB function specifically within the DG engram during memory consolidation, and identify a novel group of CREB target genes involved in the encoding of long-term memory.

Overall design Biological replicates: Fear conditioned: n=14, No shock controls: n=4, Home cage controls: n=3. The contents 10 dVenus+ and 10 dVenus- cells were aspirated from each animal (biological replicate)

Contributor(s) [Rao-Ruiz P, Couey JJ, Marcelo IM, Bouwkamp CG, Slump DE, Matos MR, van der Loo RJ, Martins GJ, van den Hout M, van IJcken WF, Costa RM, van den Oever MC, Kushner SA](#)

Citation(s) Rao-Ruiz P, Couey JJ, Marcelo IM, Bouwkamp CG et al. Engram-specific transcriptome profiling of contextual memory consolidation. *Nat Commun* 2019 May 20;10(1):2232. PMID: [31110186](#)

Submission date Mar 29, 2019  
Last update date May 23, 2019  
Contact name Steven Kushner  
E-mail(s) [s.kushner@erasmusmc.nl](mailto:s.kushner@erasmusmc.nl)  
Organization name Erasmus MC

[der Loo RJ, Martins GJ, van den Hout M, van IJcken WF, Costa RM, van den Oever MC, Kushner SA](#)

Citation(s) Rao-Ruiz P, Couey JJ, Marcelo IM, Bouwkamp CG et al. Engram-specific transcriptome profiling of contextual memory consolidation. *Nat Commun* 2019 May 20;10(1):2232. PMID: [31110186](#)

Submission date Mar 29, 2019  
Last update date May 23, 2019  
Contact name Steven Kushner  
E-mail(s) [s.kushner@erasmusmc.nl](mailto:s.kushner@erasmusmc.nl)  
Organization name Erasmus MC  
Department Psychiatry  
Street address Dr. Molewaterplein 40  
City Rotterdam  
State/province Zuid Holland  
ZIP/Postal code 3015 GD  
Country Netherlands

Platforms (1) [GPL17021](#) Illumina HiSeq 2500 (Mus musculus)

Samples (38) [GSM3690851](#) Fear Conditioned, dVenus+, MouseNr 01 [G01]  
[GSM3690852](#) Fear Conditioned, dVenus+, MouseNr 03 [G03]  
[GSM3690853](#) Fear Conditioned, dVenus-, MouseNr 01 [NG01]

**Relations**  
BioProject [PRJNA529794](#)  
SRA [SRP189843](#)



Download family	Format
<a href="#">SOFT formatted family file(s)</a>	SOFT <a href="#">?</a>
<a href="#">MINiML formatted family file(s)</a>	MINiML <a href="#">?</a>
<a href="#">Series Matrix File(s)</a>	TXT <a href="#">?</a>

Supplementary file	Size	Download	File type/resource
<a href="#">GSE129024_RAW.tar</a>	3.2 Mb	<a href="#">(http)(custom)</a>	TAR (of TXT)
<a href="#">GSE129024_counts_per_gene_sample.txt.gz</a>	647.9 Kb	<a href="#">(ftp)(http)</a>	TXT

[SRA Run Selector](#) [?](#)

Raw data are available in SRA  
Processed data provided as supplementary file  
Processed data are available on Series record

# Accessing the data through NCBI GEO

  
Gene Expression Omnibus

COVID-19 is an emerging, rapidly evolving situation.  
Get the latest public health information from CDC: <https://www.coronavirus.gov>.  
Get the latest research from NIH: <https://www.nih.gov/coronavirus>.  
Find NCBI SARS-CoV-2 literature, sequence, and clinical content: <https://www.ncbi.nlm.nih.gov/sars-cov-2/>.

[HOME](#) [SEARCH](#) [SITE MAP](#) [GEO Publications](#) [FAQ](#) [MIAME](#) [Email GEO](#)

[NCBI > GEO > Accession Display](#) [Not logged in | Login](#)

GEO help: Mouse over screen elements for information.

Scope:  Format:  Amount:  GEO accession:

**Series GSE129024** [Query DataSets for GSE129024](#)

Status Public on Mar 30, 2019  
Title Engram-specific transcriptome profiling of contextual memory consolidation  
Organism [Mus musculus](#)  
Experiment type Expression profiling by high throughput sequencing  
Summary Sparse populations of neurons in the dentate gyrus (DG) of the hippocampus are causally implicated in the encoding of contextual fear memories. However, engram-specific molecular mechanisms underlying memory consolidation remain largely unknown. Here we perform unbiased RNA sequencing of DG engram neurons 24h after contextual fear conditioning to identify transcriptome changes specific to memory consolidation. DG engram neurons exhibit a highly distinct pattern of gene expression, in which CREB-dependent transcription features prominently (P=6.2x10<sup>-13</sup>), including Atf3 (P=2.4x10<sup>-41</sup>), Penk (P=1.3x10<sup>-15</sup>), and Kcnq3 (P=3.1x10<sup>-12</sup>). Moreover, we validate the functional relevance of the RNAseq findings by establishing the causal requirement of intact CREB function specifically within the DG engram during memory consolidation, and identify a novel group of CREB target genes involved in the encoding of long-term memory.

Overall design Biological replicates: Fear conditioned: n=14, No shock controls: n=4, Home cage controls: n=3. The contents 10 dVenus+ and 10 dVenus- cells were aspirated from each animal (biological replicate)

Contributor(s) [Rao-Ruiz P, Couey JJ, Marcelo IM, Bouwkamp CG, Slump DE, Matos MR, van der Loo RJ, Martins GJ, van den Hout M, van IJcken WF, Costa RM, van den Oever MC, Kushner SA](#)

Citation(s) Rao-Ruiz P, Couey JJ, Marcelo IM, Bouwkamp CG et al. Engram-specific transcriptome profiling of contextual memory consolidation. *Nat Commun* 2019 May 20;10(1):2232. PMID: [31110186](#)

Submission date Mar 29, 2019  
Last update date May 23, 2019  
Contact name Steven Kushner  
E-mail(s) [s.kushner@erasmusmc.nl](mailto:s.kushner@erasmusmc.nl)  
Organization name Erasmus MC

[der Loo RJ, Martins GJ, van den Hout M, van IJcken WF, Costa RM, van den Oever MC, Kushner SA](#)

Citation(s) Rao-Ruiz P, Couey JJ, Marcelo IM, Bouwkamp CG et al. Engram-specific transcriptome profiling of contextual memory consolidation. *Nat Commun* 2019 May 20;10(1):2232. PMID: [31110186](#)

Submission date Mar 29, 2019  
Last update date May 23, 2019  
Contact name Steven Kushner  
E-mail(s) [s.kushner@erasmusmc.nl](mailto:s.kushner@erasmusmc.nl)  
Organization name Erasmus MC  
Department Psychiatry  
Street address Dr. Molewaterplein 40  
City Rotterdam  
State/province Zuid Holland  
ZIP/Postal code 3015 GD  
Country Netherlands

Platforms (1) [GPL17021](#) Illumina HiSeq 2500 (Mus musculus)

Samples (38) [GSM3690851](#) Fear Conditioned, dVenus+, MouseNr 01 [G01]  
[GSM3690852](#) Fear Conditioned, dVenus+, MouseNr 03 [G03]  
[GSM3690853](#) Fear Conditioned, dVenus-, MouseNr 01 [NG01]






**Relations**  
BioProject [PRJNA529794](#)  
SRA [SRP189843](#)

Download family	Format
<a href="#">SOFT formatted family file(s)</a>	SOFT <a href="#">?</a>
<a href="#">MINiML formatted family file(s)</a>	MINiML <a href="#">?</a>
<a href="#">Series Matrix File(s)</a>	TXT <a href="#">?</a>

Supplementary file	Size	Download	File type/resource
<a href="#">GSE129024_RAW.tar</a>	3.2 Mb	<a href="#">(http)</a> <a href="#">(custom)</a>	TAR (of TXT)
<a href="#">GSE129024_counts_per_gene_sample.txt.gz</a>	647.9 Kb	<a href="#">(ftp)</a> <a href="#">(http)</a>	TXT


[SRA Run Selector](#) [?](#)  
Raw data are available in SRA  
Processed data provided as supplementary file  
Processed data are available on Series record

# Accessing the metadata

 **SRA Run Selector**     Log in to NIH

**Filters List**




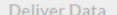

- 1 ☐ AvgSpotLen
- 2 ☐ Bases
- 3 ☐ Bytes
- 4 ☐ Cell\_type
- 5 ☐ source\_name
- 6 ☐ Treatment

Accession   **Search**



**Common Fields**










LibraryLayout	SINGLE
LibrarySelection	CDNA
LibrarySource	TRANSCRIPTOMIC
Organism	Mus musculus
Platform	ILLUMINA
ReleaseDate	2019-04-01
SRA Study	<a href="#">SRP189843</a>
Strain	C57BL/6J
Tissue	Hippocampal Dentate Gyrus granule neurons

**Select**

	Runs	Bytes	Bases	Download	Cloud Data Delivery	Computing
<b>Total</b>	38	23.92 Gb	46.00 G	Metadata  Accession List		
<b>Selected</b>	0	0	0	Metadata  Accession List  JWT Cart		

**Found 38 Items**

<input checked="" type="checkbox"/>			<b>Run</b> <sup>1</sup>		<b>BioSample</b> <sup>2</sup>		<b>AvgSpotLen</b> <sup>3</sup>		<b>Bases</b> <sup>4</sup>		<b>Bytes</b> <sup>5</sup>	<b>Cell_type</b> <sup>6</sup>		<b>Experiment</b> <sup>7</sup>	<b>GEO_Accession</b> <sup>8</sup>		<b>Mouse_ID</b> <sup>9</sup>		<b>Sample Name</b> <sup>10</sup>	<b>SampleID</b> <sup>11</sup>	<b>source_name</b> <sup>12</sup>	<b>Treatment</b> <sup>13</sup>
-------------------------------------	---	---	-------------------------	---	-------------------------------	---	--------------------------------	---	---------------------------	---	---------------------------	-------------------------------	---	--------------------------------	-----------------------------------	---	------------------------------	---	----------------------------------	-------------------------------	----------------------------------	--------------------------------

# Accessing the metadata

NCBI SRA Run Selector 🔍 ? ⚙️ 🔗 Log in to NIH

Accession  🔍 Search

**Filters List**

- ☐ AvgSpotLen
- ☐ Bases
- ☐ Bytes
- ☐ Cell\_type
- ☐ source\_name
- ☐ Treatment

**Common Fields**

LibraryLayout	SINGLE
LibrarySelection	CDNA
LibrarySource	TRANSCRIPTOMIC
Organism	Mus musculus
Platform	ILLUMINA
ReleaseDate	2019-04-01
SRA Study	<a href="#">SRP189843</a>
Strain	C57BL/6J
Tissue	Hippocampal Dentate Gyrus granule neurons

**Select**

	Runs	Bytes	Bases	Download	Cloud Data Delivery	Computing
Total	38	23.92 Gb	46.00 G	<span>Metadata</span> <span>or</span> <span>Accession List</span>		
Selected	0	0	0	<span>Metadata</span> <span>or</span> <span>Accession List</span> <span>or</span> <span>JWT Cart</span>	<span>Deliver Data</span>	<span>Galaxy</span>

**Found 38 Items**  🔍 Clear

<input checked="" type="checkbox"/>	<input type="checkbox"/>	Run	BioSample	AvgSpotLen	Bases	Bytes	Cell_type	Experiment	GEO_Accession	Mouse_ID	Sample Name	SampleID	source_name	Treatment
-------------------------------------	--------------------------	-----	-----------	------------	-------	-------	-----------	------------	---------------	----------	-------------	----------	-------------	-----------