Cours du 28 novembre 2011

Human brain evolution: harnessing the genomics (r)evolution to link genes, cognition, and behavior

Konopka G, Geschwind DH

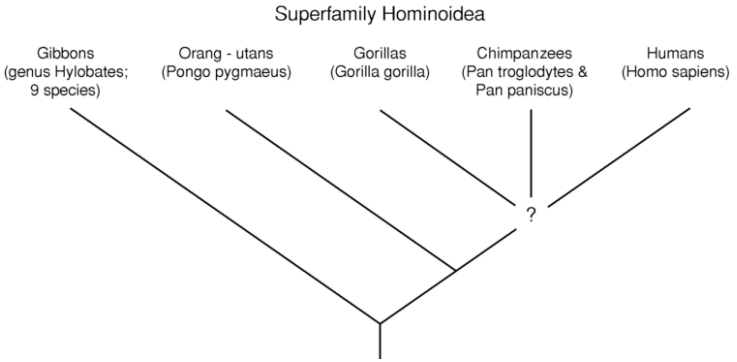
Neuron
2010 vol. 68 (2) pp. 231–44

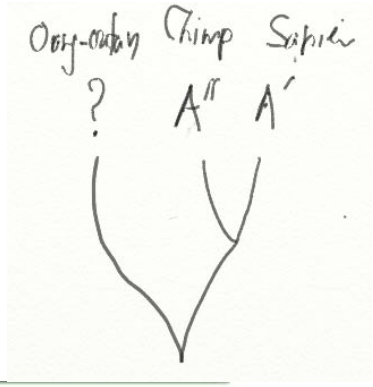
Studying the brain from an evolutionary perspective and combining these results with those from development and pathology, connecting genetic variation to neural circuit development and functioning, will yield the best approximation of how natural forces shaped this organ. Aside from satisfying basic curiosity about the origins of our abilities, such endeavors have enormous implications for understanding human diseases involving cognition and behavior, ranging from intellectual disability and autism to neurodegenerative dementias.

Human brain evolution: harnessing the genomics (r)evolution to link genes, cognition, and behavior

Konopka G, Geschwind DH

Neuron
2010 vol. 68 (2) pp. 231-44





Common Name	Species	Original Date Sequenced	Unique Features
Mouse	*Mus musculus*	2002	Common model system for neurobehavioral/ neurogenetic experiments
Human	*Homo sapiens*	2001	
Rat	*Rattus norvegicus*	2003	Common model system for neurobehavioral experiments; now available for genetic manipulations
Fly	*Drosophila melanogaster*	2003	Common model system for neurobehavioral/ neurogenetic experiments
Worm	*Caenorhabditis elegans*	2004	Aging and longevity
Chimpanzee	*Pan troglodytes*	2005	Great ape; most similar to human on genome level
Zebrafish	*Danio rerio*	2005	Easy to visualize brain development; behavior; genetically malleable
Rhesus macaque	*Macaca mulatta*	2006	Old world monkey; bred in the USA for behavior and neurophysiology experiments
Honey bee	*Apis mellifera*	2006	Social behavior; aggression
Orangutan	*Pongo pygmaeus abelii*	2007	Great ape; useful for outgroup comparison with human/chimp
Elephant shark	*Callorhinchus milii*	2007	Outgroup for zebrafish and mouse comparisons
Dolphin	*Tursiops truncatus*	2008	Evidence for vocalization and self-awareness
Zebra finch	*Taeniopygia guttata*	2008	Patterned vocalization
Sea hare	*Aplysia californica*	2008	Learning and memory
Elephant	*Loxodonta africana*	2009	Evidence for vocalization and self-awareness
Pig	*Sus scrofa*	2009	Gyrencephalic cortex
Bat	*Myotis lucifugus*	2010	Echolocation
Ferret	*Mustela putorius furo*	in progress	Gyrencephalic cortex
Cichlids	*Tilapia nilotica*	in progress	"Natural" mutants for the study of evolution
Marmoset	*Callithrix jacchus*	in progress	New world monkey

Human brain evolution: harnessing the genomics (r)evolution to link genes, cognition, and behavior

Konopka G, Geschwind DH

Neuron
2010 vol. 68 (2) pp. 231–44

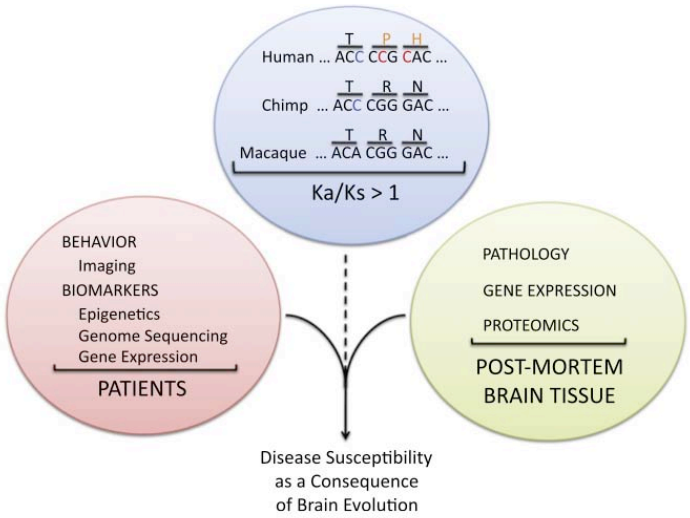


Figure 1. Combining Comparative Genomics with Phenotype and Expression Data Leads to Disease Insights

Measures of positive selection are typically calculated by dividing the number of nonsynonymous changes (depicted in red) by the number of synonymous changes (depicted in blue); a value greater than one is used as evidence for positive selection. It should be emphasized that this is arbitrary and only takes into account known protein coding regions. Screening genomes for genes under positive selection is one important step; however, other measures such as links to behavior and expression need to be incorporated. Furthermore, a gene does not have to have undergone positive selection to be a disease-susceptibility gene. Other changes in evolution such as timing or location of expression could make a gene or signaling pathway vulnerable in disease.

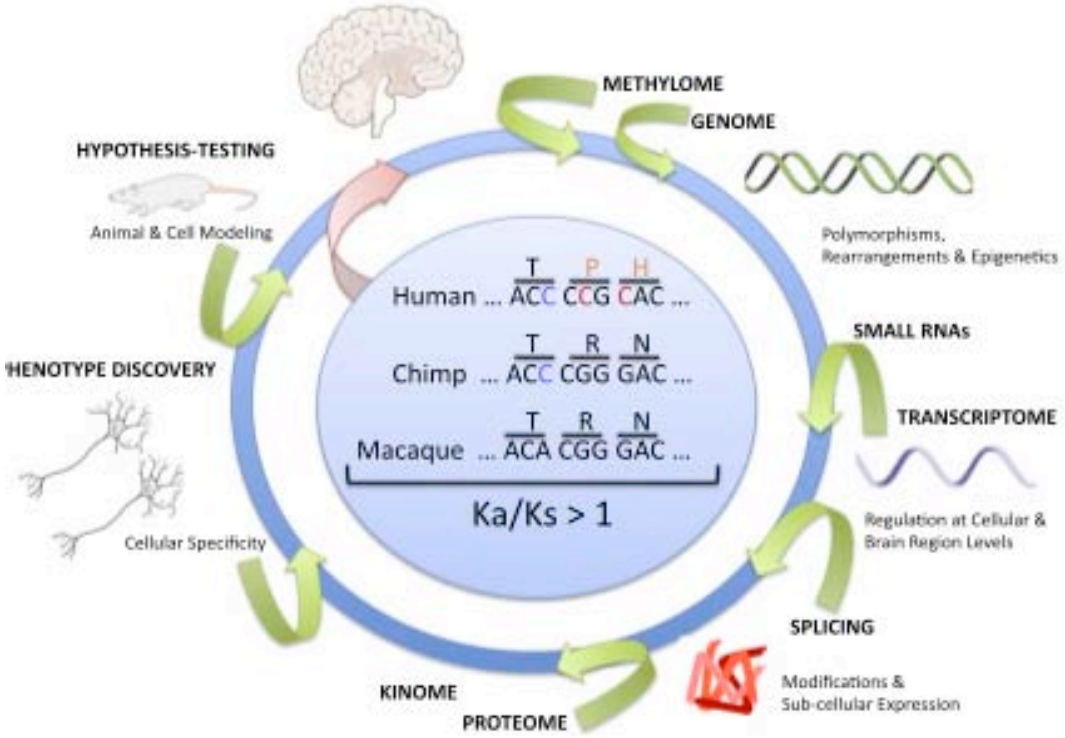


Figure 2. Multiple Layers of Regulation Underlie Human Brain Evolution

More than genomic comparisons need to be considered when building our understanding of human brain evolution. Regulation at the level of the epigenome (e.g., methylome), regulation of expression by changes in transcription or small RNA regulation, novel isoforms through differential splicing, changes in regional and subcellular expression, and posttranslational modifications (e.g., kinome) all need to be taken into account, and many of these changes can be queried using NGS techniques. Also, the incorporation of animal models and cell-specific lines of inquiry need to undertaken. Finally, the integration of all of these data will lead to phenotype discovery and hypothesis-testing that ultimately will inform polymorphism discovery in human brain diseases. In this manner, modern neurogenetic investigations are not only an exercise in advanced data analysis, but comprehensive data integration.

Both noncoding and protein-coding RNAs contribute to gene expression evolution in the primate brain

Genome Biol Evol
2010 vol. 2 pp. 67–79

Babbitt CC, Fedrigo O, Pfefferle AD, Boyle AP, Horvath JE, Furey TS, Wray G

Categorical Enrichment To determine functional category enrichment for the differentially expressed genes, we employed the PANTHER (HMM Library Version 6.0; Mi et al. 2005) and GO (The Gene Ontology Consortium, 2000) gene ontology databases. Our background set of genes were those genes measured in our tissue samples. PANTHER and GO category enrichment scores were computed using the top 5% of the hypergeometric probability distribution. Python code used to perform all enrichments is available at: http://www.duke.edu/~ofedrigo/Olivier_Fedrigo/PythonScripts.html.

Categorical Enrichments for Differentially Expressed Genes between the Human and Chimpanzee Individuals

Category	*P* value	Top 5.0%	Total
PANTHER			
DNA repair	8.06×10^{-05}	17	119
DNA metabolism	9.58×10^{-05}	26	233
Intracellular protein traffic	0.0001132	61	759
Electron transport	0.007361	16	163
Neurotransmitter release	0.01398	10	90
Oxidative phosphorylation	0.01562	7	53
Induction of apoptosis	0.01983	10	95
Endocytosis	0.02428	17	202
Extracellular transport and import	0.03121	6	48
Protein targeting and localization	0.03184	14	162
Nuclear transport	0.03957	7	64
Cytokinesis	0.04562	7	66
GO			
Translational elongation	0.0004988	11	68
Viral genome replication	0.002211	4	12
Protein import into nucleus, docking	0.008714	4	17
Phospholipid metabolic process	0.01311	4	19
Transport	0.0139	34	460
Glutamate signaling pathway	0.01946	3	12
Induction of apoptosis	0.02269	10	97
Intracellular protein transport	0.02412	14	156
tRNA aminoacylation for protein translation	0.02438	3	13
Lipid catabolic process	0.02472	7	58
Nucleocytoplasmic transport	0.0299	3	14
Regulation of GTPase activity	0.03603	3	15
Electron transport chain	0.04029	8	78
RNA processing	0.04054	6	51
Base-excision repair	0.04274	3	16
Inactivation of MAPK activity	0.04274	3	16
Protein stabilization	0.04274	3	16
DNA repair	0.04792	11	125
Phospholipid biosynthetic process	0.04874	4	28

NOTE.—The results for the biological process domain of both the GO and PANTHER ontologies are shown. Categorical enrichments are for the top 5% of a hypergeometric probability distribution. The right-hand columns show the number of genes in the top 5%, as well as the total number of genes evaluated. Categories that evaluated less than 10 genes total are not shown. Categories are further colored according to hierarchically related ontology terms: nucleic acid metabolism (green), electron transport (yellow), neuronal activity (blue), transport, extra- and intracellular protein traffic (pink), and lipid metabolism (purple).

Contrasts between adaptive coding and noncoding changes during human evolution

Haygood R, Babbitt CC, Fedrigo O, Wray G

Proc Natl Acad Sci USA
2010 vol. 107 (17) pp. 7853-7

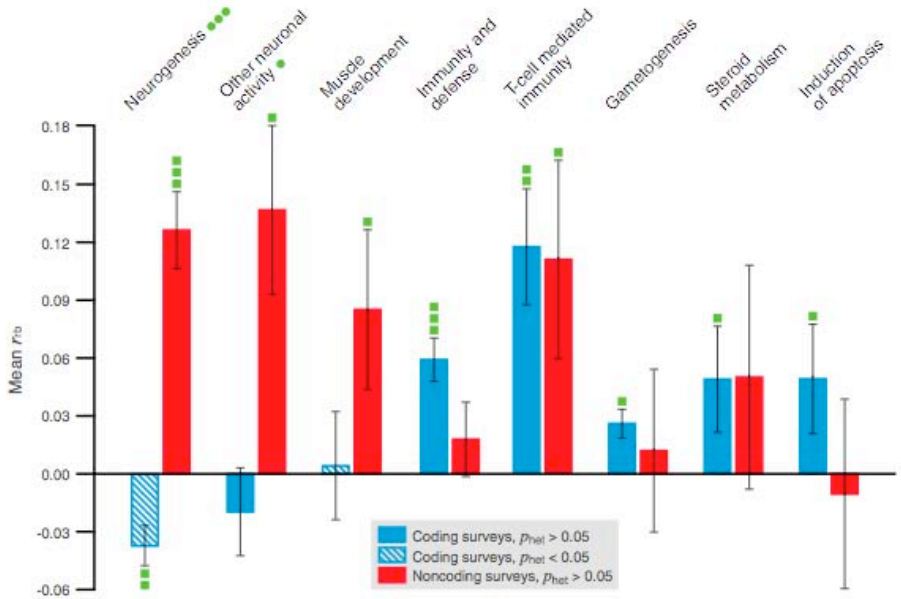


Fig. 1. Large PANTHER biological processes enriched with positive selection across coding or noncoding surveys. Plotted categories are all but three of those for which (*i*) the mean number of genes per survey is at least 50 across all surveys, (*ii*) the mean rank-biserial correlation (r_{rb}) between score for positive selection and membership in the category is significantly positive ($P_{err} < 0.05$) across coding or noncoding surveys, and (*iii*) heterogeneity is nonsignificant ($P_{het} > 0.05$) across the same surveys. The three such categories not plotted are "ectoderm development," "neuronal activities," and "mesoderm development," which are not significantly enriched when their respective subcategories "neurogenesis," "other neuronal activity," and "muscle development" are subtracted. Error bars represent SEM- r_{rb}. Green blocks on error bars indicate one-tailed *P* values for the one-sample *z* test of mean $r_{rb} = 0$. One block, $5 \times 10^{-4} < P < 0.05$; two blocks, $5 \times 10^{-6} < P < 5 \times 10^{-4}$; three blocks, $P < 5 \times 10^{-6}$. Similarly, green dots on category names indicate two-tailed *P* values for two-sample *z* tests of equal mean, r_{rb}, across coding versus noncoding surveys.

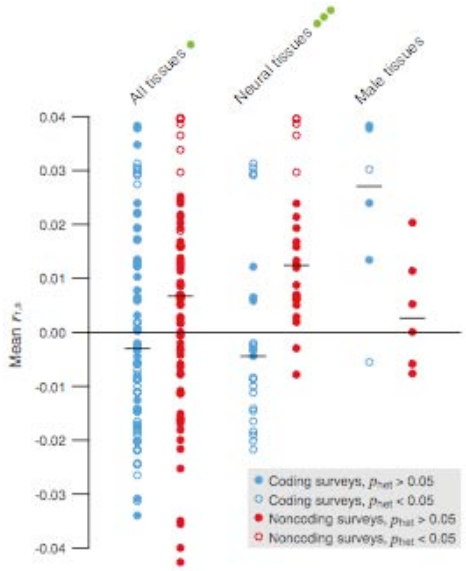


Fig. 3. Expression specificity and positive selection across coding and noncoding surveys. Neural tissues and male (reproductive) tissues are highlighted. Circles represent mean rank correlation ($r_{r,s}$) between the score for positive selection and specificity to tissues in the Novartis Gene Expression Atlas. Horizontal lines indicate median of means. Green dots on group names indicate two-tailed *P* values for the Wilcoxon signed-rank test of equal median across coding versus noncoding surveys: one dot, $5 \times 10^{-4} < P < 0.05$; three dots, $P < 5 \times 10^{-6}$.

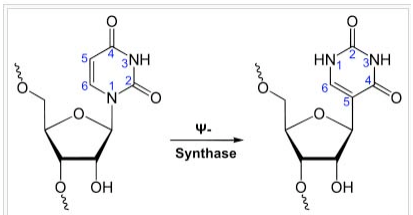
Table 1 | **Types of ncRNAs***

Name	Size	Location	Number in humans	Functions	Illustrative examples	Refs
Short ncRNAs						
miRNAs	19–24 bp	Encoded at widespread locations	>1,424	Targeting of mRNAs and many others	miR-15/16, miR-124a, miR-34b/c, miR-200	3–8
piRNAs	26–31bp	Clusters, intragenic	23,439	Transposon repression, DNA methylation	piRNAs targeting *RASGRF1* and LINE1 and IAP elements	13–19
tiRNAs	17–18bp	Downstream of TSSs	>5,000	Regulation of transcription?	Associated with the *CAP1* gene	37
Mid-size ncRNAs						
snoRNAs	60–300 bp	Intronic	>300	rRNA modifications	U50, SNORD	20–22
PASRs	22–200 bp	5′ regions of protein-coding genes	>10,000	Unknown	Half of protein-coding genes	10
TSSa-RNAs	20–90 bp	–250 and +50 bp of TSSs	>10,000	Maintenance of transcription?	Associated with *RNF12* and *CCDC52* genes	35
PROMPTs	<200 bp	–205 bp and –5 kb of TSSs	Unknown	Activation of transcription?	Associated with *EXT1* and *RBM39* genes	36
Long ncRNAs						
lincRNAs	>200 bp	Widespread loci	>1,000	Examples include scaffold DNA–chromatin complexes	*HOTAIR*, *HOTTIP*, *lincRNA-p21*	2,28–30
T-UCRs	>200 bp	Widespread loci	>350	Regulation of miRNA and mRNA levels?	uc.283+, uc.338, uc160+	31–34
Other lncRNAs	>200 bp	Widespread loci	>3,000	Examples include X-chromosome inactivation, telomere regulation, imprinting	*XIST*, *TSIX*, TERRAs, *p15AS*, *H19*, *HYMAI*	2,23–25

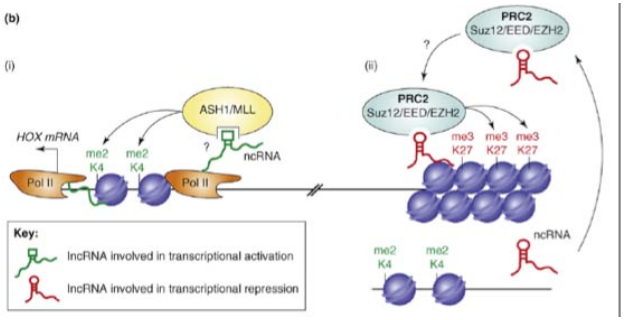
Non-coding RNAs in human disease

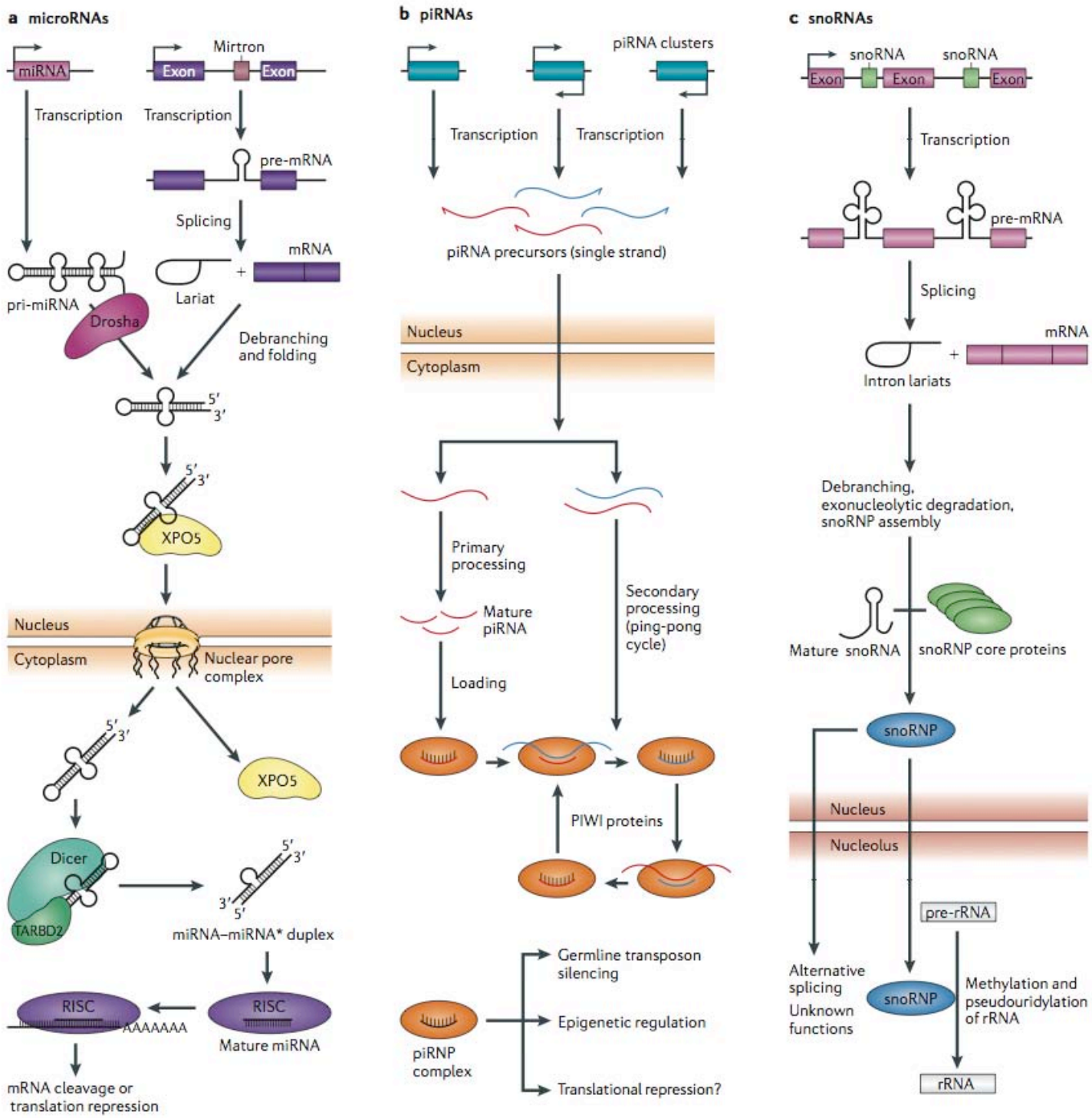
Esteller M

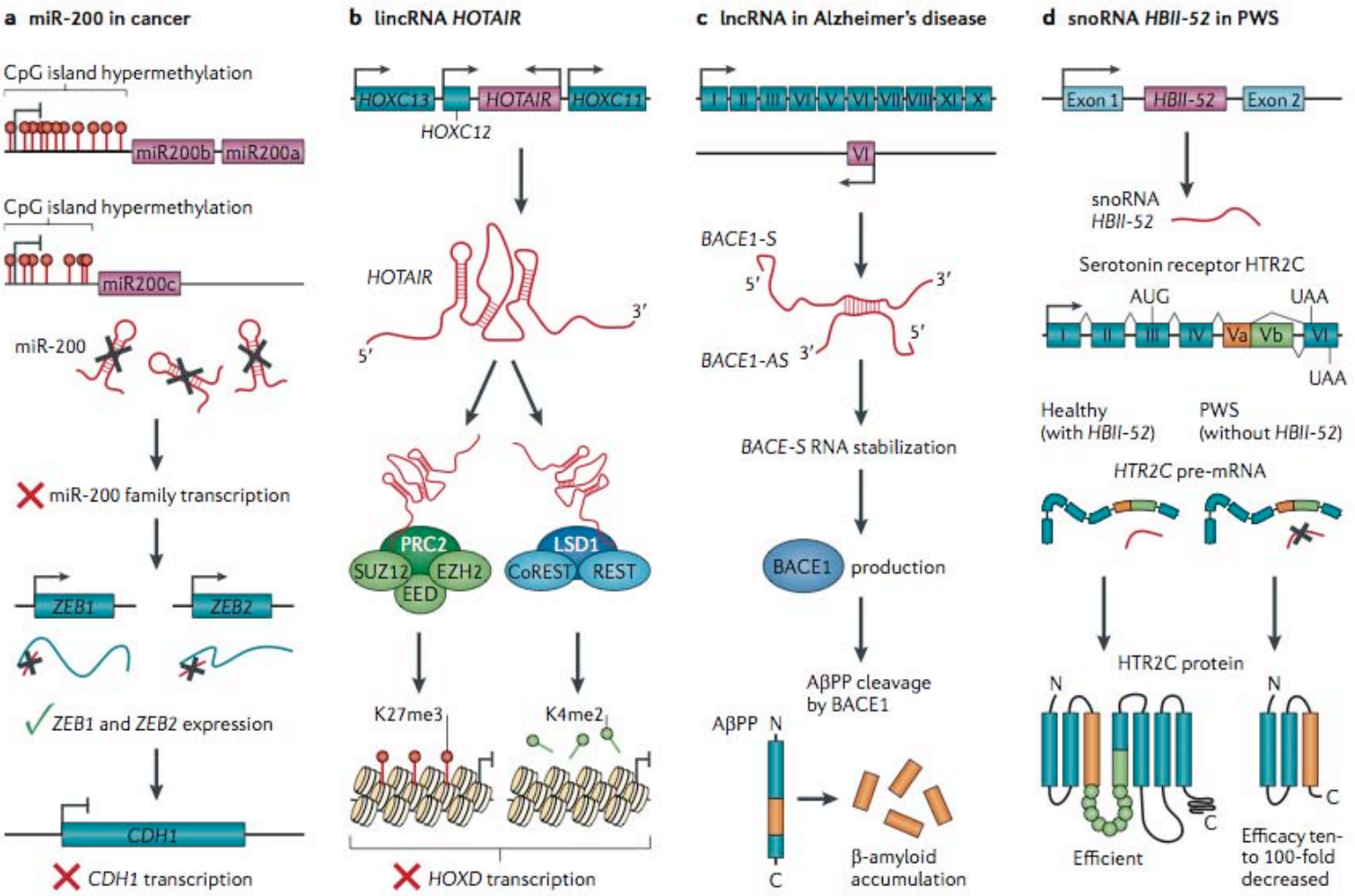
Nature Reviews Genetics 12, 861 (2011). doi: 10.1038/nrg3074



Pseudouridine is biosynthesized from uridine via the action of Ψ synthases.




a microRNAs
miRNA
Mirtron
Exon
Exon
Transcription
Transcription
pre-mRNA
Splicing
mRNA
pri-miRNA
Drosha
Lariat
Debranching and folding
5′ 3′
XPO5
Nucleus
Cytoplasm
Nuclear pore complex
XPO5
Dicer
TARBD2
miRNA–miRNA* duplex
RISC
AAAAAAA
RISC
Mature miRNA
mRNA cleavage or translation repression
b piRNAs
piRNA clusters
Transcription
Transcription
piRNA precursors (single strand)
Nucleus
Cytoplasm
Primary processing
Mature piRNA
Loading
Secondary processing (ping-pong cycle)
PIWI proteins
piRNP complex
Germline transposon silencing
Epigenetic regulation
Translational repression?
c snoRNAs
snoRNA
snoRNA
Exon
Exon
Exon
Transcription
pre-mRNA
Splicing
mRNA
Intron lariats
Debranching, exonucleolytic degradation, snoRNP assembly
Mature snoRNA
snoRNP core proteins
snoRNP
Nucleus
Nucleolus
pre-rRNA
Alternative splicing
Unknown functions
snoRNP
Methylation and pseudouridylation of rRNA
rRNA

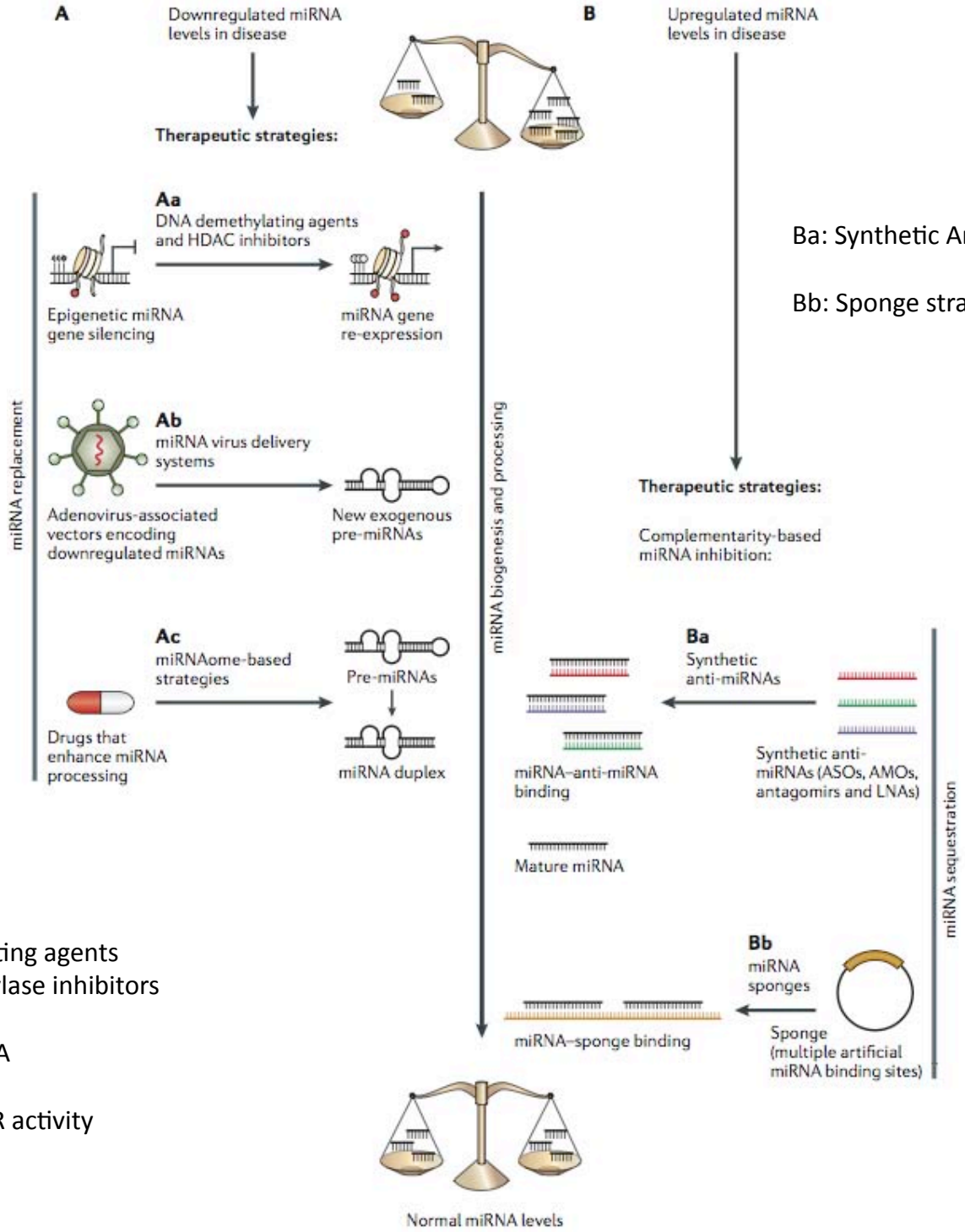


invasiveness. **c** | lncRNA targeting of β-secretase 1 (BACE1) has a role in the pathophysiology of Alzheimer's disease. An antisense lncRNA, *BACE1-AS*, regulates the expression of the sense *BACE1* gene (labelled *BACE1-S* in the figure) through the stabilization of its mRNA. *BACE1-AS* is elevated in Alzheimer's disease, increasing the amount of BACE1 protein and, subsequently, the production of β-amyloid peptide. **d** | The role of the snoRNA in Prader–Willi syndrome (PWS). The loss of the snoRNA in PWS changes the alternative splicing of the serotonin receptor *HTR2C* precursor mRNA (pre-mRNA), resulting in a protein with reduced function. AβPP, amyloid-β precursor protein; CoREST, REST corepressor.

Table 3 | **Illustrative list of ncRNAs that are disrupted in non-tumoural disorders**

Disease	Involved ncRNAs	ncRNA type	Refs
Spinal motor neuron disease	miR-9	miRNA	87
Spinocerebellar ataxia type 1	miR-19, miR-101, miR-100	miRNA	88
Amyotropic lateral sclerosis	miR-206	miRNA	86
Arrhytmia and hypertension	miR-1	miRNA	98
Atheromatosis and atherosclerosis	miR-10a, miR-145, mR-143 and miR-126	miRNA	100–102
Atheromatosis and atherosclerosis	Circular ncRNA linked to the CDKN2A locus	lncRNA	119
Cardiac hypertrophy	miR-21	miRNA	144
Rett's syndrome	miR-146a, miR-146b, miR-29 and miR-382	miRNA	108,109
5q syndrome	miR-145 and miR-146a	miRNA	106
ICF syndrome	miR-34b, miR-34c, miR-99b, let-7e and miR-125a	miRNA	107
Crohn's disease	miR-196	miRNA	110
Prader–Willi and Angelman syndromes	snoRNA cluster at 15q11–q13 imprinted locus	snoRNA	114–116
Beckwith–Wiedeman syndrome	lncRNAs *H19* and *KCNQ1OT1*	lncRNA	145
Uniparental disomy 14	snoRNA cluster at 14q32.2 imprinted locus	snoRNA	145
Silver–Russell syndrome	lncRNA *H19*	lncRNA	145
Silver–Russell syndrome	miR-675	miRNA	145
McCune–Albright syndrome	lncRNA *NESP-AS*	lncRNA	145
Deafness	miR-96	miRNA	111
Alzheimer's disease	miR-29, miR-146 and miR-107	miRNA	89–91
Alzheimer's disease	ncRNA antisense transcript for *BACE1*	lncRNA	112
Parkinson's disease	miR-7, miR-184 and let-7	miRNA	82
Down's syndrome	mir-155 and miR-802	miRNA	83
Idiopathic neurodevelopmental disease	T-UCRs uc.195, uc.392, uc.46 and uc.222	T-UCR	113
Rheumatoid arthritis	miR-146a	miRNA	147
Transient neonatal diabetes mellitus	lncRNA *HYMAI*	lncRNA	148
Pseudohypoparathyroidism	lncRNA *NESP-AS*	lncRNA	146

BACE1, β-secretase 1; CDKN2A, cyclin-dependent kinase inhibitor 2A; *HYMAI*, hydatidiform mole associated and imprinted; ICF syndrome, immunodeficiency, centromeric region instability and facial anomalies syndrome; KCNQ1OT1, KCNQ1 opposite strand/antisense transcript 1; lncRNA, long non-coding; miRNA, microRNA; NESP, also known as GNAS; NESP-AS, NESP antisense; ncRNA, non-coding RNA; snoRNA, small nucleolar RNA; T-UCR, transcribed ultraconserved region.



Ba: Synthetic Antisens strategy

Bb: Sponge strategy

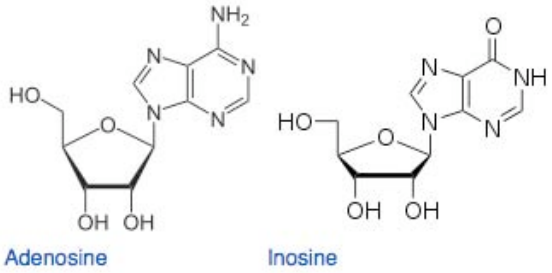
Aa: DNA demethylating agents and histone deacetylase inhibitors

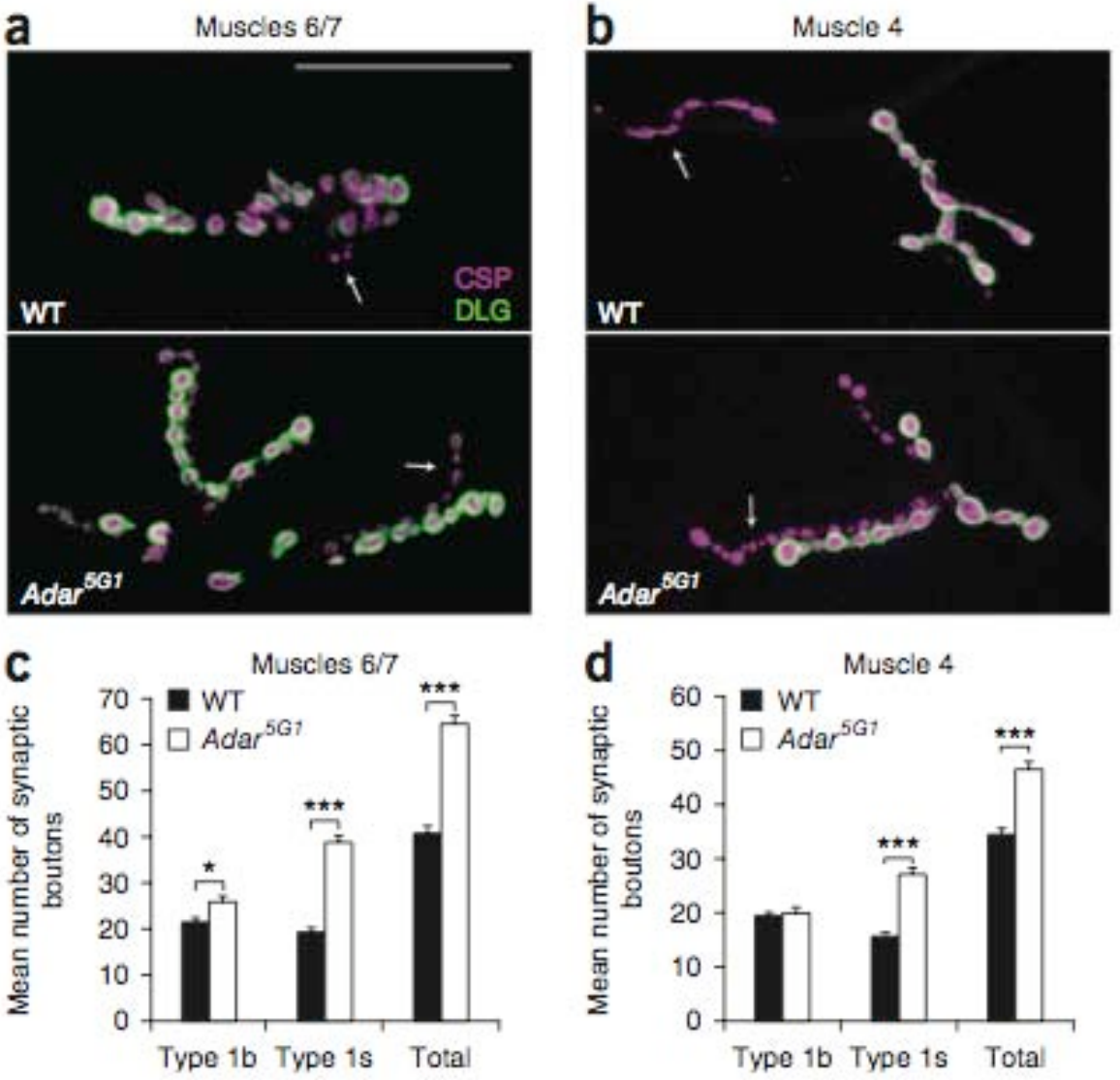
Ab: Delivery of µRNA

Ac: Enhancing DICER activity

Modulation of dADAR-dependent RNA editing by the Drosophila fragile X mental retardation protein

Bhogal B, Jepson JE, Savva YA, Pepper AS, Reenan RA, Jongens TA



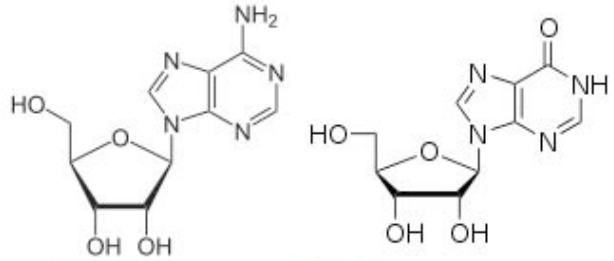


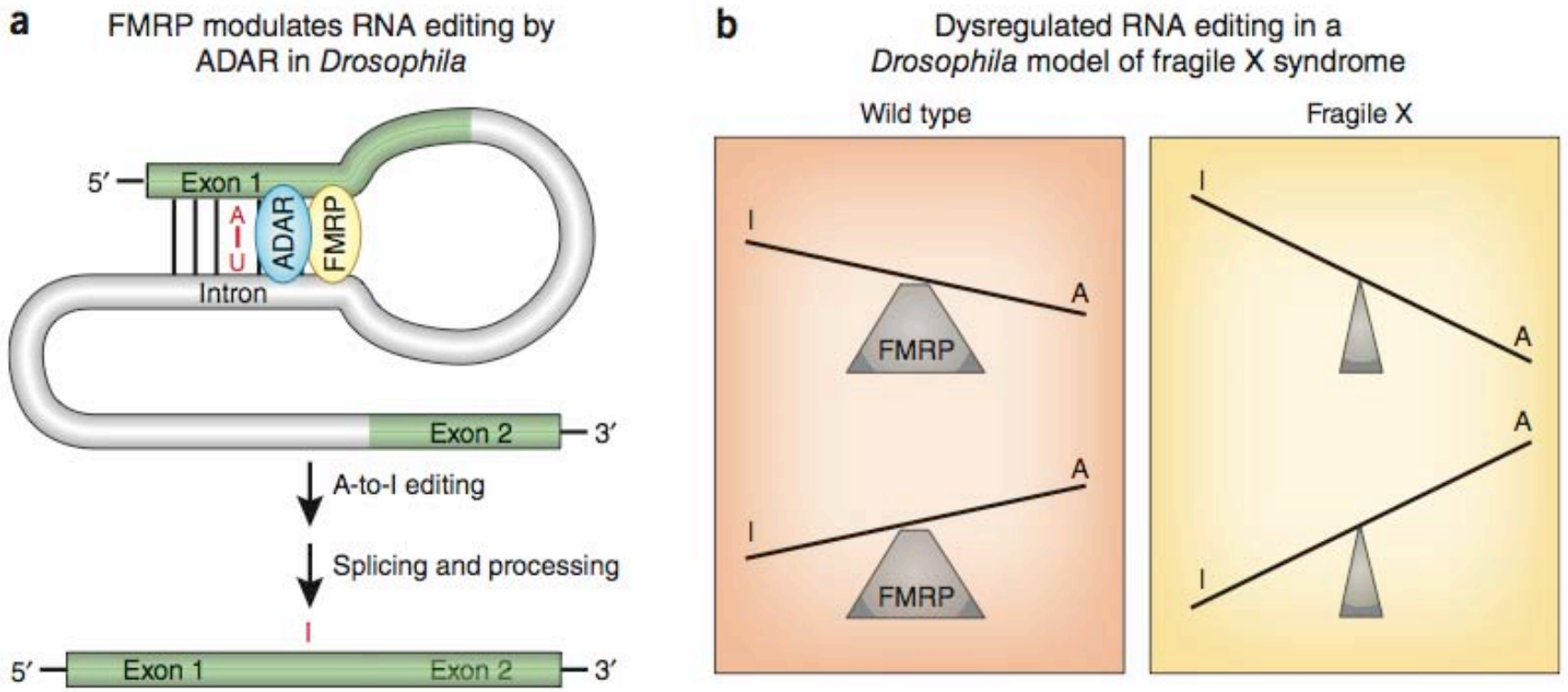
Fragile balance: RNA editing tunes the synapse

Bassell G

Nature Neuroscience 14, 1492 (2011). doi: 10.1038/nn.2982

Nature Neuroscience
2011 vol. 14 (12) pp. 1492–1494





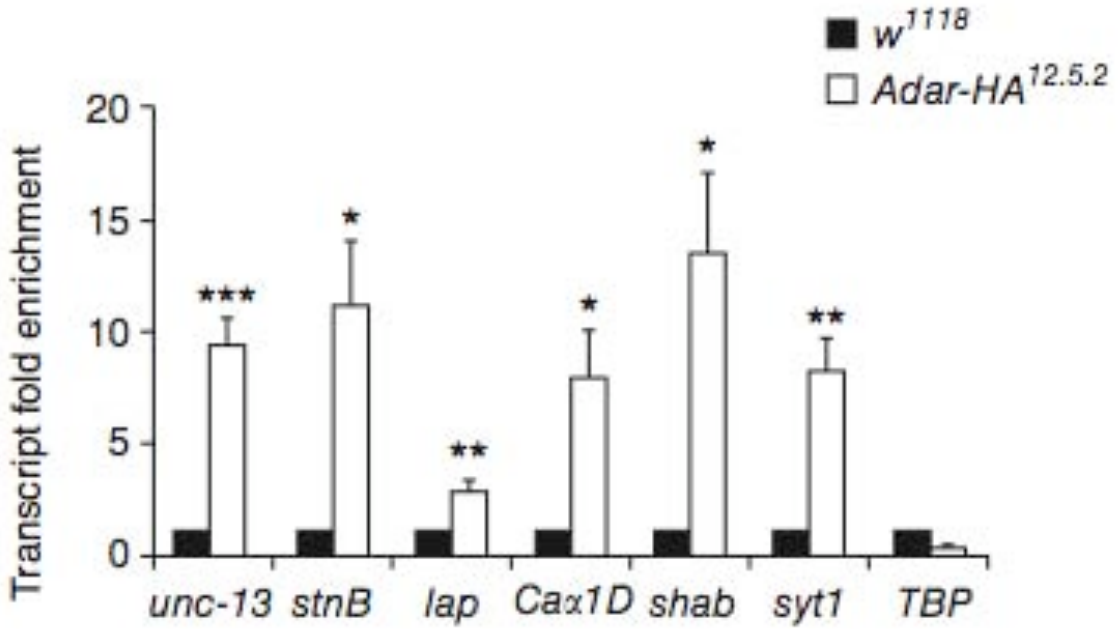
Katie Vicari

Figure 1 Regulation of A-to-I RNA editing by FMRP in *Drosophila*. (**a**) The data of Bhogal *et al.*[2], together with other studies, suggest a model in which ADAR and FMRP cooperate in the nucleus to regulate ADAR activity, possibly on double-stranded RNA targets comprising exonic and intronic sequences. (**b**) In wild-type neurons (left), the amount of RNA editing is tightly controlled or fine-tuned by FMRP, whereas RNA editing is imbalanced and dysregulated in fragile X syndrome (right), as a result of a loss of control at the fulcrum.

Modulation of dADAR-dependent RNA editing by the Drosophila fragile X mental retardation protein

Bhogal B, Jepson JE, Savva YA, Pepper AS, Reenan RA, Jongens TA

NATURE NEUROSCIENCE VOLUME 14 | NUMBER 12 | DECEMBER 2011



unc-13: libération des neuromédiateurs,

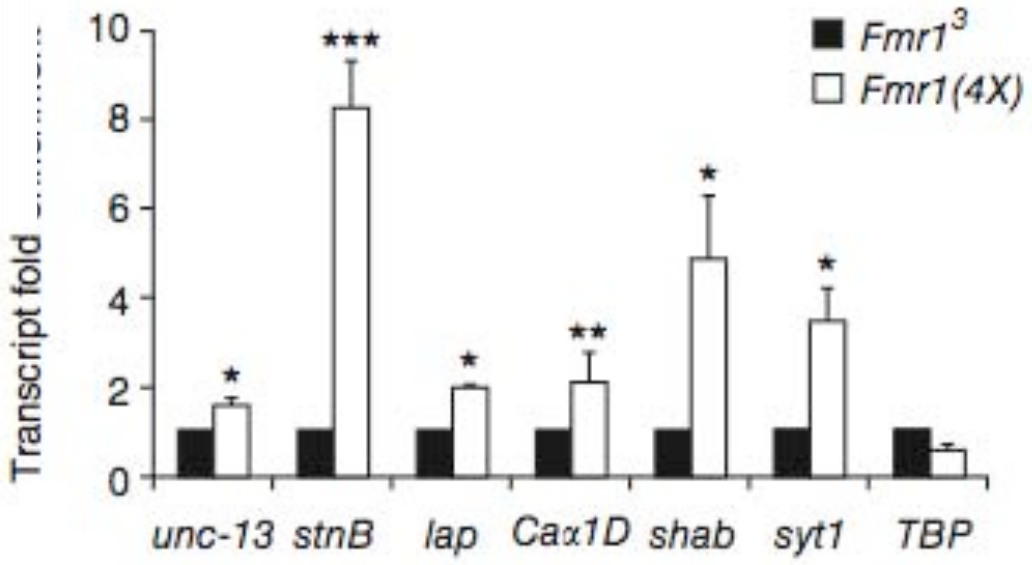
stnb: transmission synaptique, endocytose de vésicules synaptique,

lap et syt1: fonctions inconnues

Ca-alpha1D: canal calcique

Shab: canal potassique

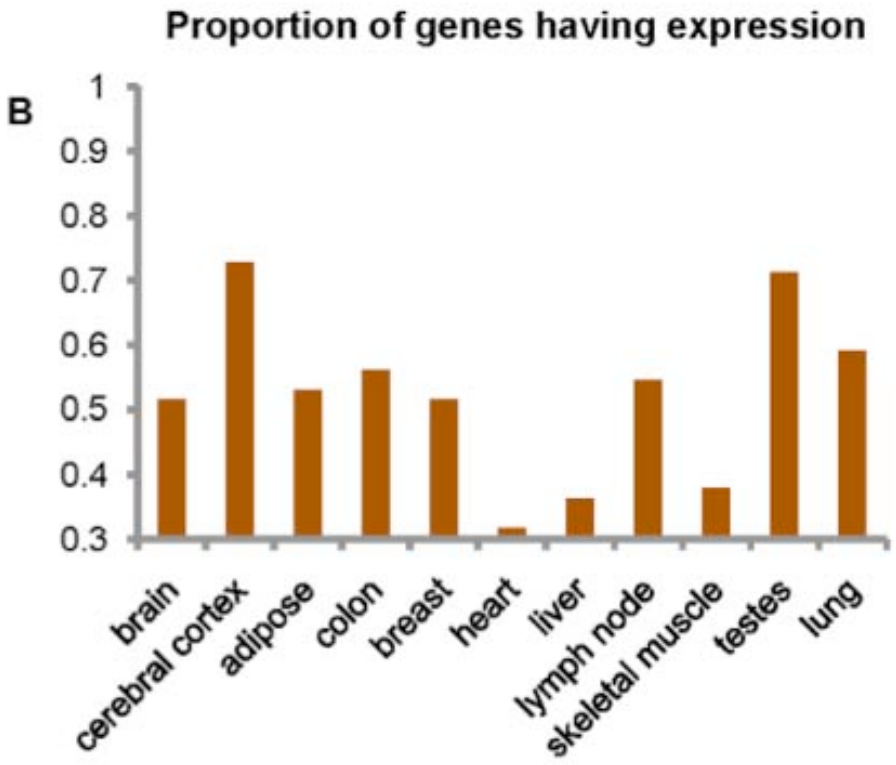
TBP: initiation de la transcription

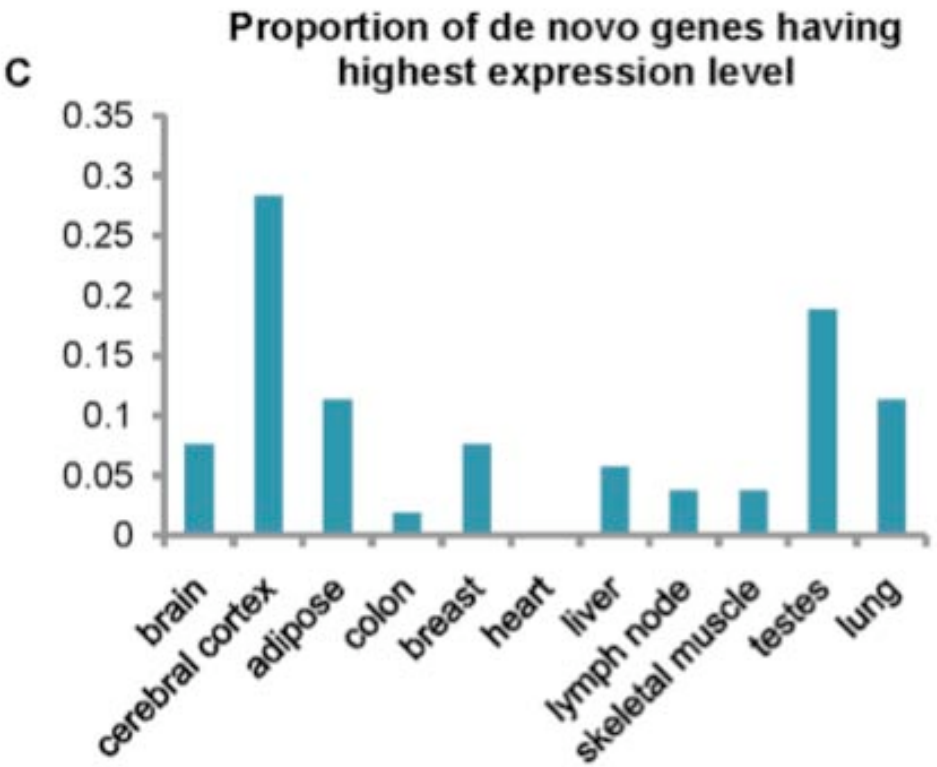


De Novo Origin of Human Protein-Coding Genes

Wu D, Irwin DM, Zhang Y

PLoS Genet
2011 vol. 7 (11) pp. e1002379

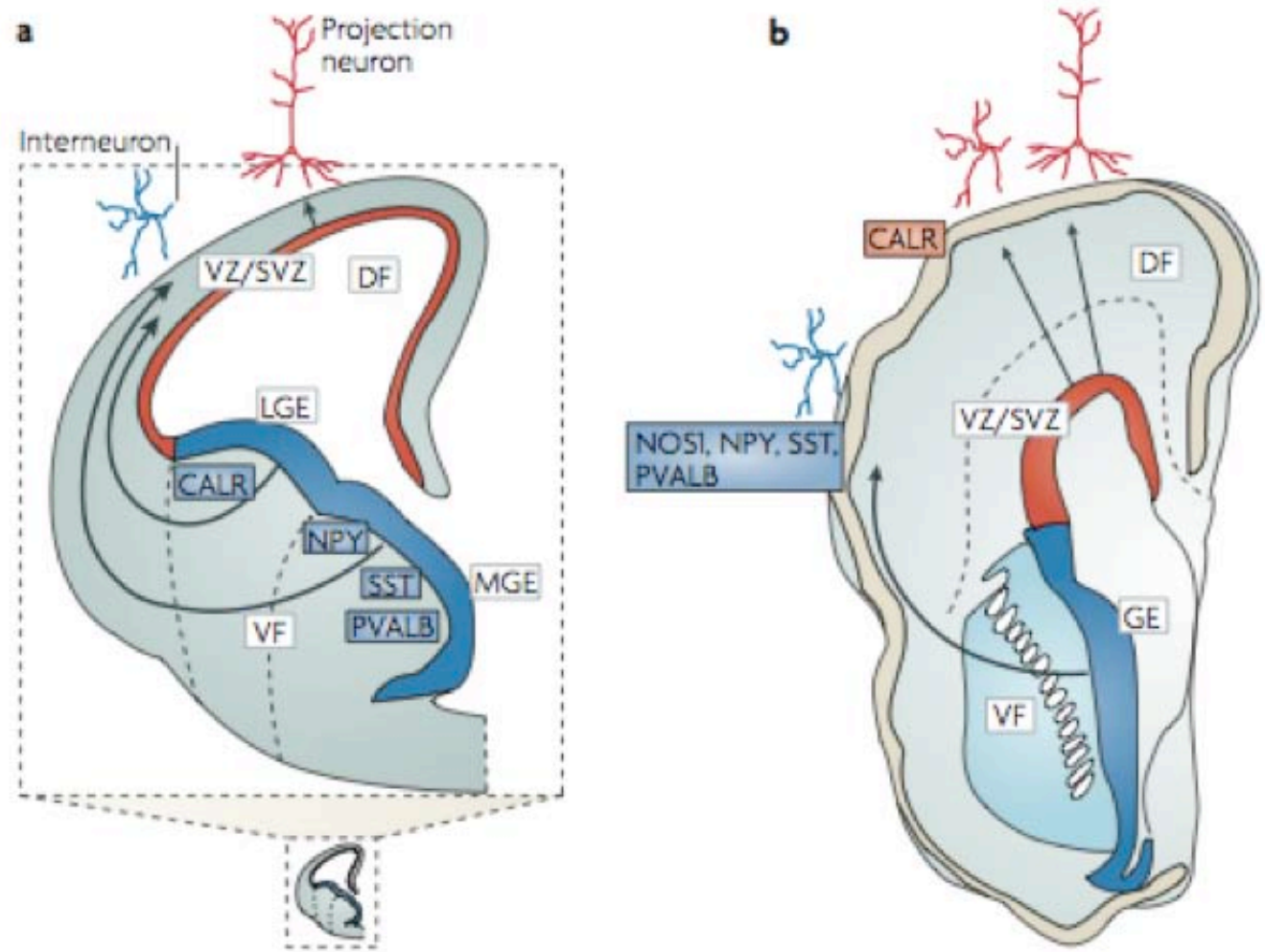




Evolution of the neocortex: a perspective from developmental biology

Rakic P

Nat Rev Neurosci
2009 vol. 10 (10) pp. 724-35

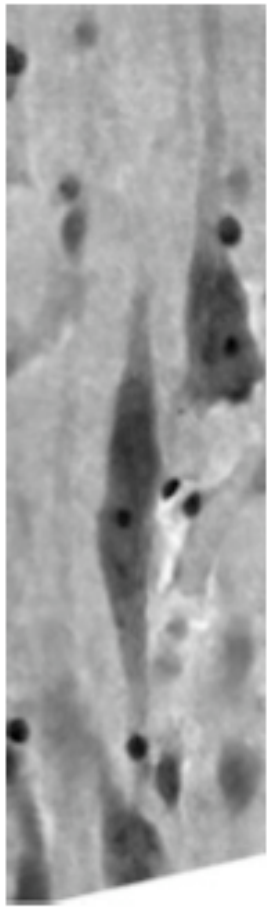


The von Economo neurons in the frontoinsular and anterior cingulate cortex

Allman JM, Tetreault NA, Hakeem AY, Manaye KF, Semendeferi K, Erwin JM, Park S, Goubert V, Hof PR

Division of Biology, California Institute of Technology, Pasadena, California, USA.

Ann N Y Acad Sci
2011 vol. 1225 pp. 59-71



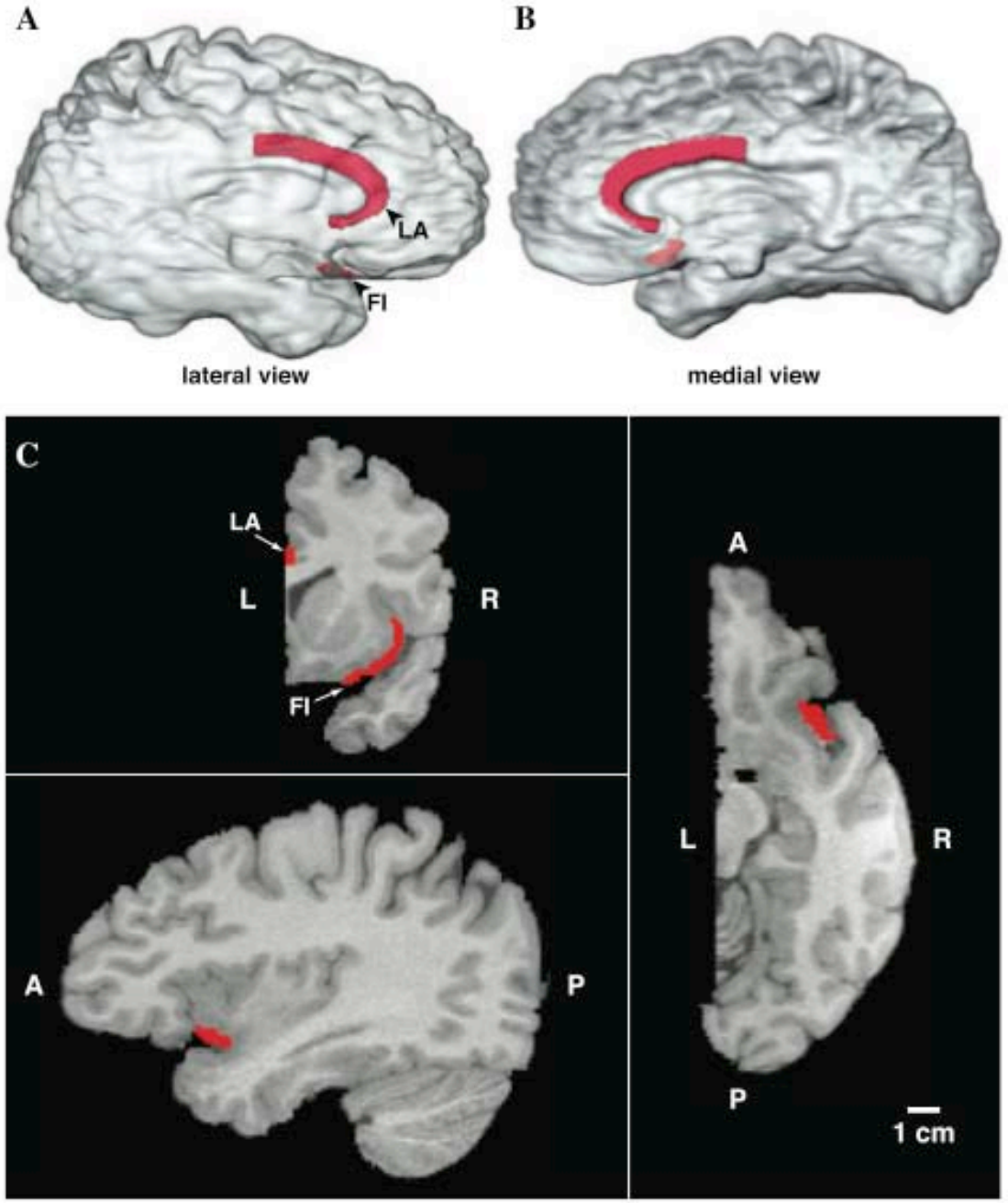


Figure 1. The location of the VEN-containing areas FI and LA indicated (in red) on the MR scan images of the right hemisphere of a young adult human female. A and B are three-dimensional reconstructions of the hemisphere. C shows frontal (upper left), parasagittal (lower left), and horizontal (right) slices through the hemisphere.

The von Economo neurons in the frontoinsular and anterior cingulate cortex

Allman JM, Tetreault NA, Hakeem AY, Manaye KF, Semendeferi K, Erwin JM, Park S, Goubert V, Hof PR

Division of Biology, California Institute of Technology, Pasadena, California, USA.

Ann N Y Acad Sci
2011 vol. 1225 pp. 59–71

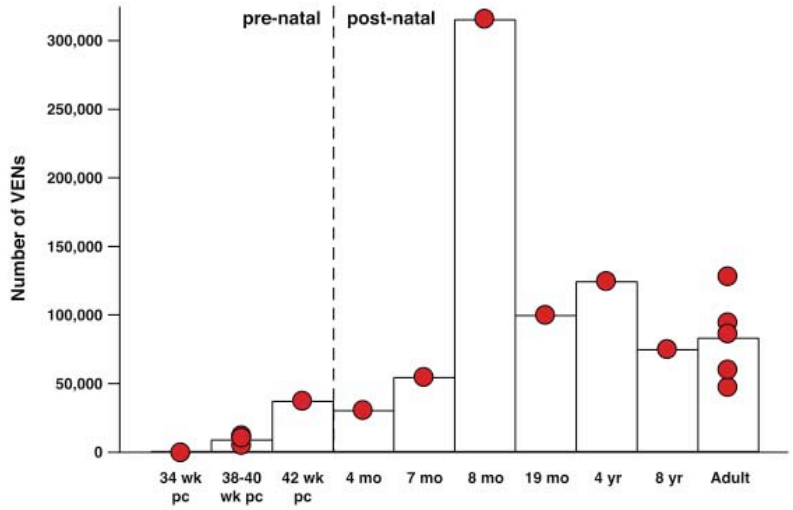


Figure 5. The number of VENs increases after birth. The number of VENs in right hemisphere FI in humans of different ages. VEN numbers are low in neonates and increase after birth. The 8-month-old individual examined had markedly more VENs in the right hemisphere than any other subject in this study; this might possibly be due to individual variation. The right hemisphere VEN measurement in this individual was repeated with similar results. The difference between the number of VENs in right FI for pre- and postnatal subjects was statistically significant ($P = 0.0029$), and this significance remained when the 8-month-old individual was removed from the comparison ($P = 0.0040$). The number of VENS in left FI and in both hemispheres together was also significantly different for pre- and postnatal individuals ($P = 0.0056$ for both). Significance was determined using the Mann–Whitney U-test.

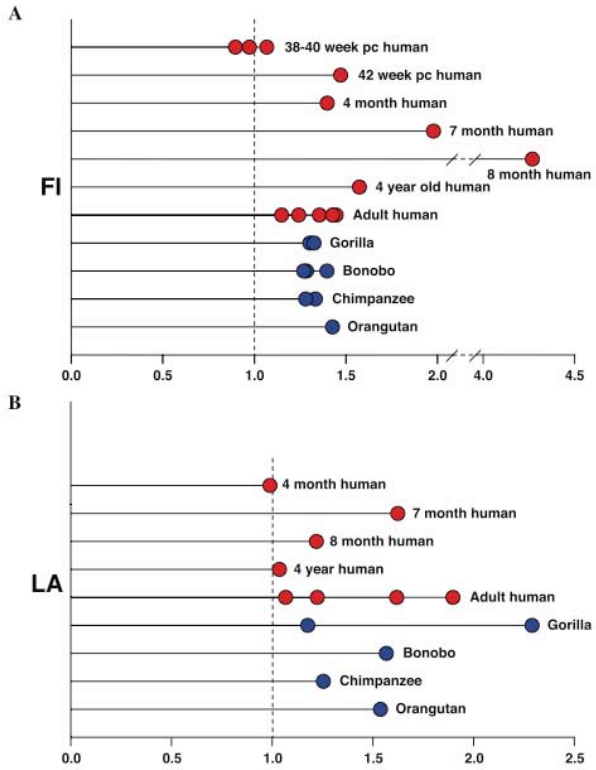


Figure 6. The ratio of the number of VENs in the right hemisphere to the number of VENs in the left hemisphere. (A) In postnatal humans and great apes, there are consistently more VENs in FI on the right side. This ratio develops after birth. In neonates, the numbers in each hemisphere are almost even, while in infants, juveniles, and adults there are many more VENs in the right hemisphere. When the numbers of VENs in the right and left hemispheres were compared for FI in the postnatal cases, the difference was statistically significant both with and without the 8-month-old outlier ($P = 0.0039$ for all postnatal humans and $P = 0.0078$ without the 8-month-old case). For postnatal apes and humans combined, the hemispheric difference for FI was significant at $P < 0.0001$. (B) The ratio of VENs in right and left LA. This ratio is less consistent than in area FI, but in almost all cases there are more VENs on the right side. When the number of VENs in the right and left hemispheres in postnatal humans was compared for LA, the result was statistically significant ($P = 0.03$). When postnatal apes and humans were combined, the difference was significant at $P = 0.001$. Significance was determined using the Mann–Whitney U-test.