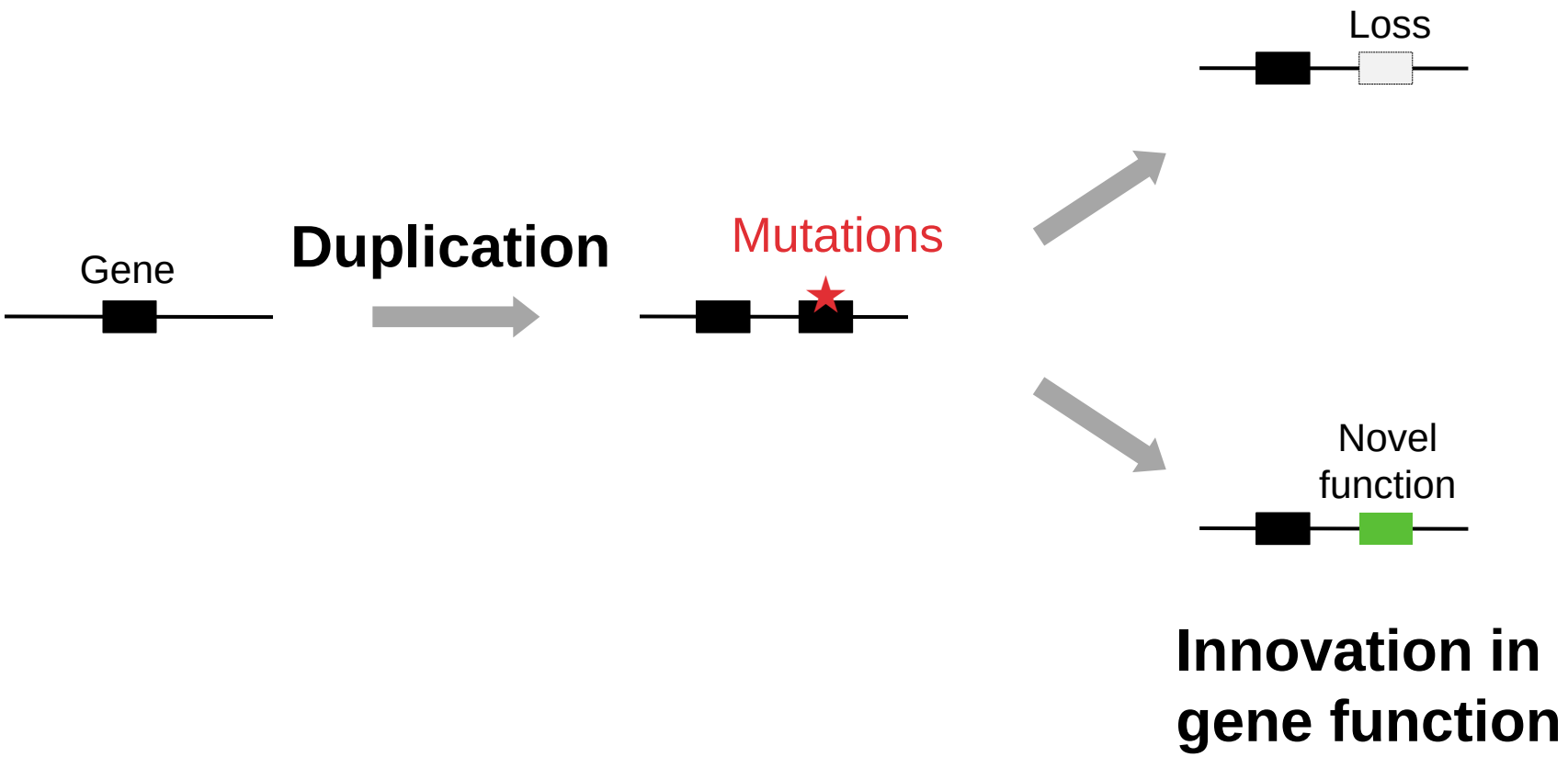
Duplicated genes maintained after whole genome duplications and their relation to human disorders

Takashi Makino

Graduate School of Life Sciences
Tohoku University

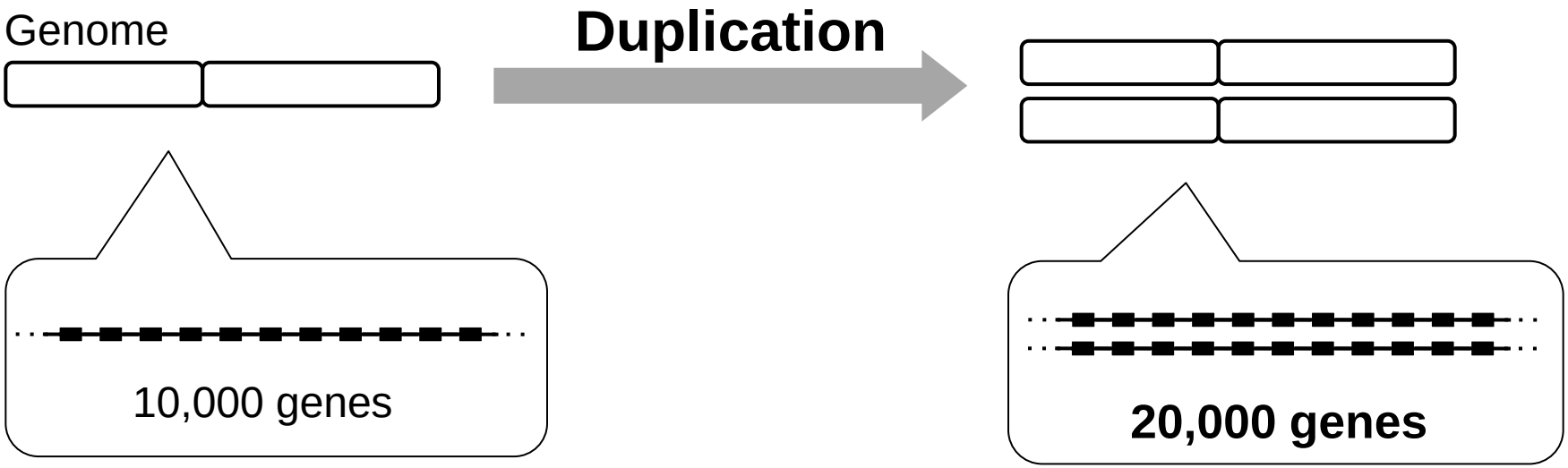
Gene duplication



Innovation in gene function

Small-**S**cale gene **D**uplication **(SSD)**

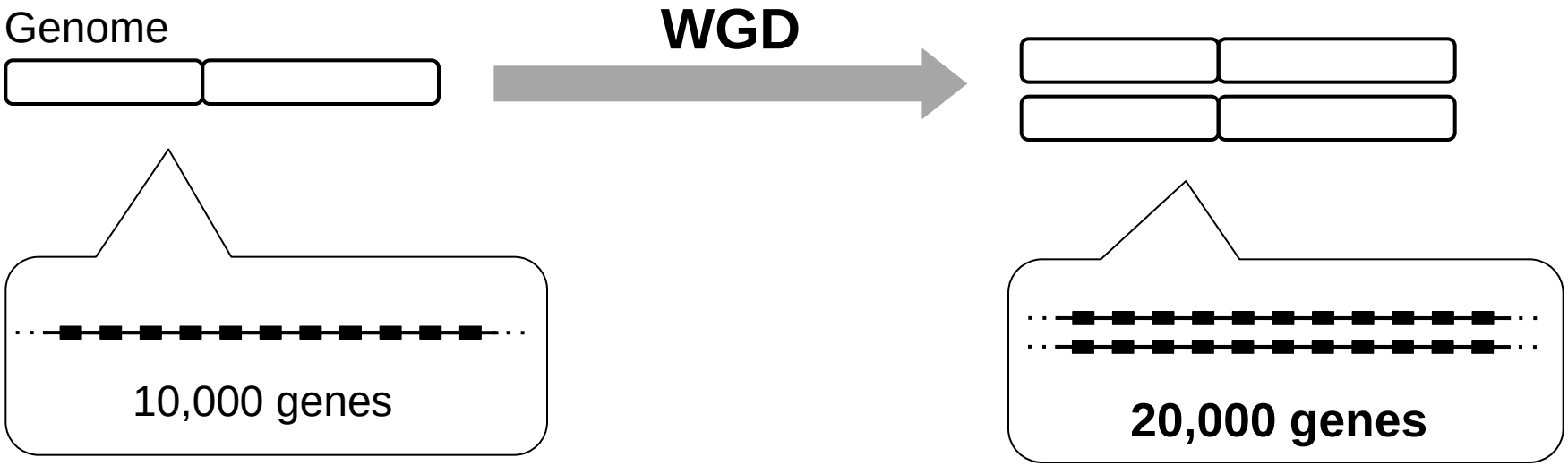
Whole **G**enome **D**uplication **(WGD)**



Dose change by duplication

Gene dosage
1

Gene dosage
2



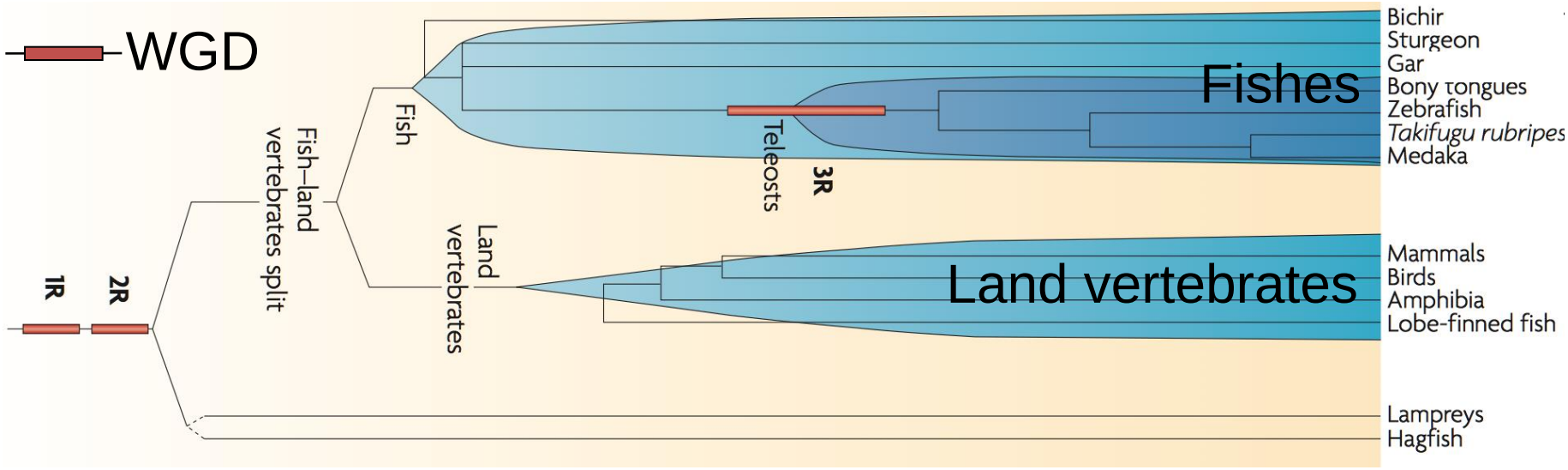
WGD increases gene dosage for all genes.

WGD in vertebrate linage

Dr. Susumu Ohno

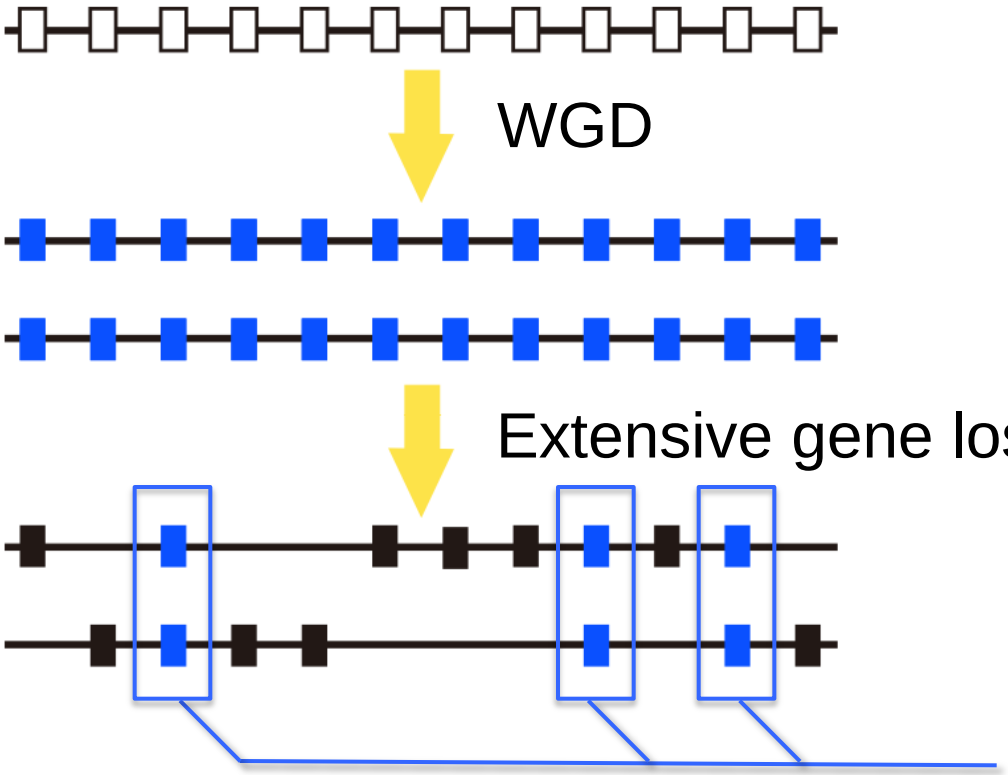
2R hypothesis

Two rounds of genome duplication in the early vertebrate lineage



Peer and Meyer, Nat. Rev. Gen. (2010)

Extensive gene losses after WGD



Makino and McLysaght.
Genome Research (2012)

30% of human genes can be traced back to 2R WGD events.

Random gene retention after WGD? → **No**

Ohnologs

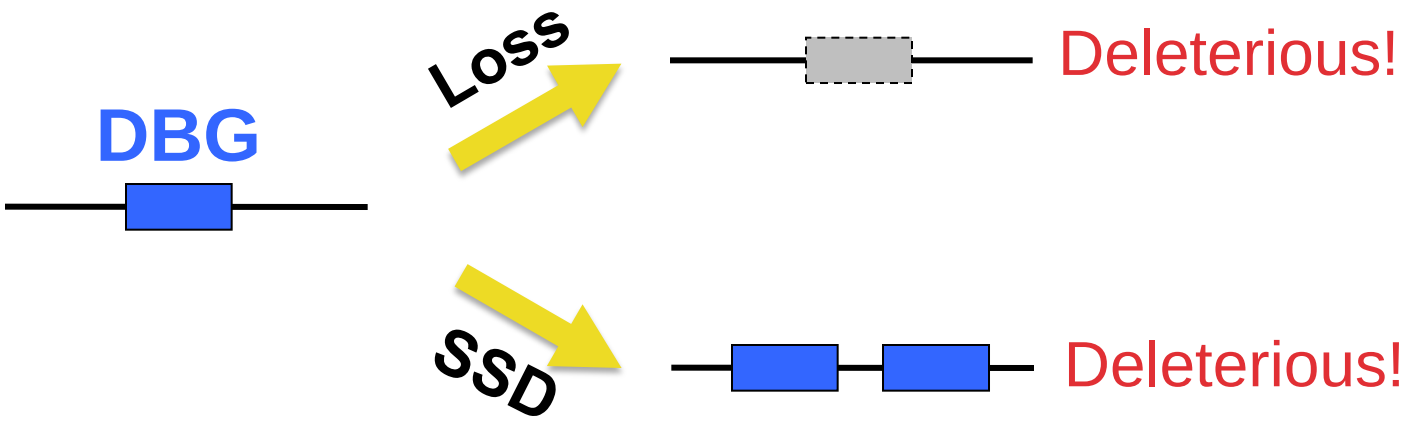
Functional bias

Developmental genes, Transcription factors,
Protein complex members

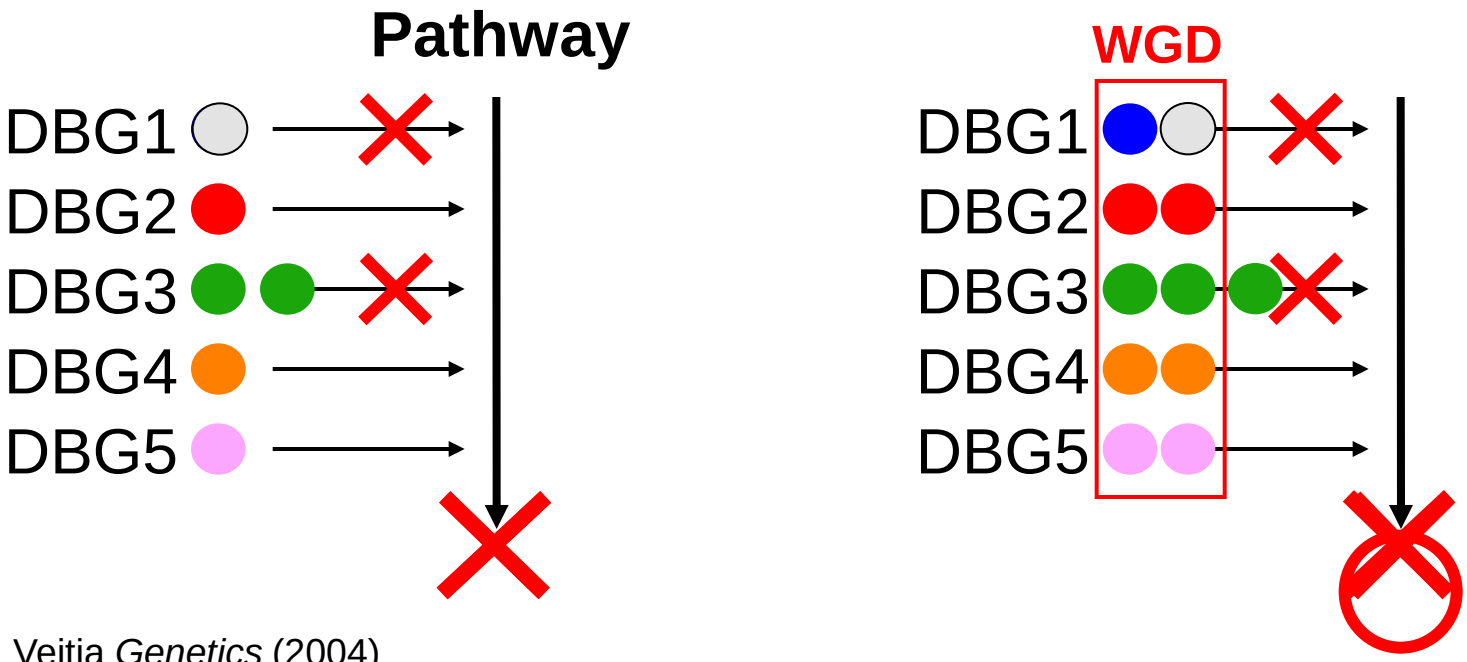
Makino *et al. Trends Genet* (2009), Maere *et al. PNAS* (2005), Blomme *et al. Gen Biol* (2006)

Dosage-balanced genes (DBGs)

DBGs are not robust to gene loss and small-scale gene duplication (SSD). Veitia. Bioessays (2002) and Papp *et al.* Nature (2003)



Dosage balance hypothesis



Veitia *Genetics* (2004)

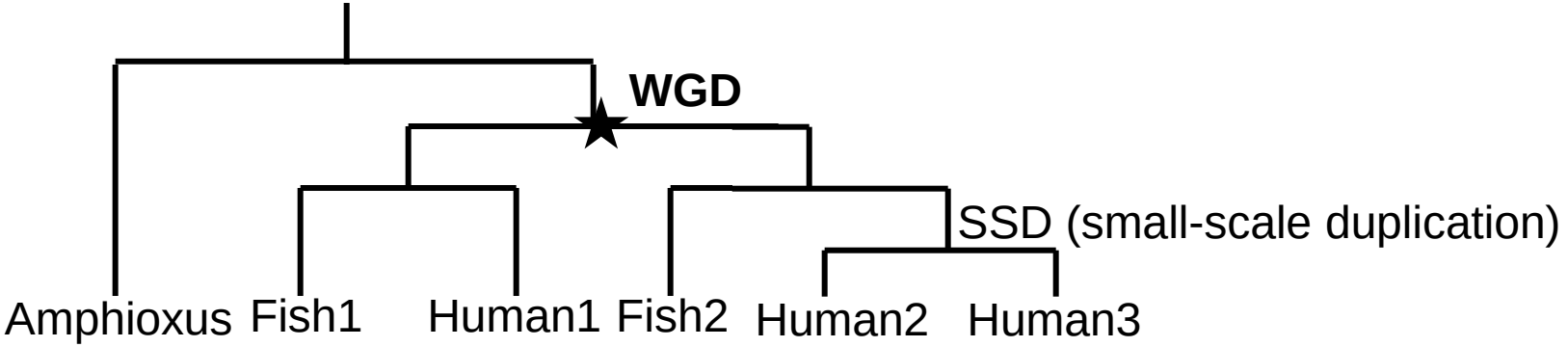
WGD duplicates all genes simultaneously and therefore does not perturb relative dosages.

Identification of human ohnologs

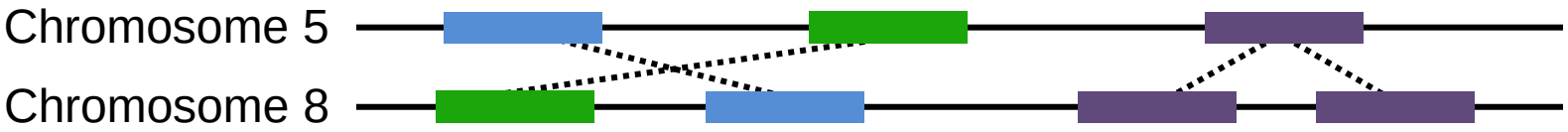
1. Homology

Homology search by blastp (human, fishes and amphioxus)
(E-value < e-7; alignable > 30%)

2. Timing of gene duplication events



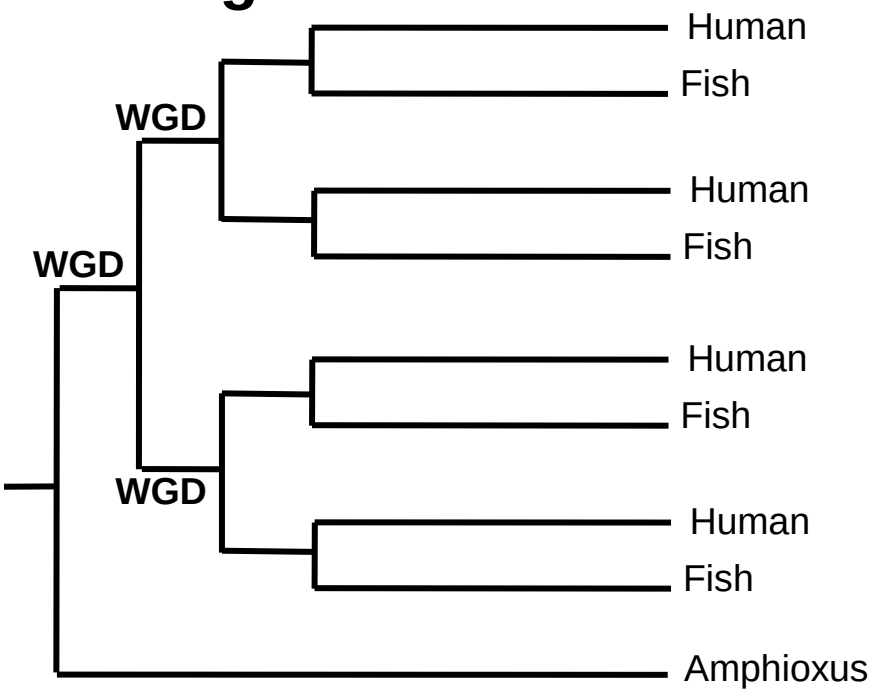
3. Gene synteny



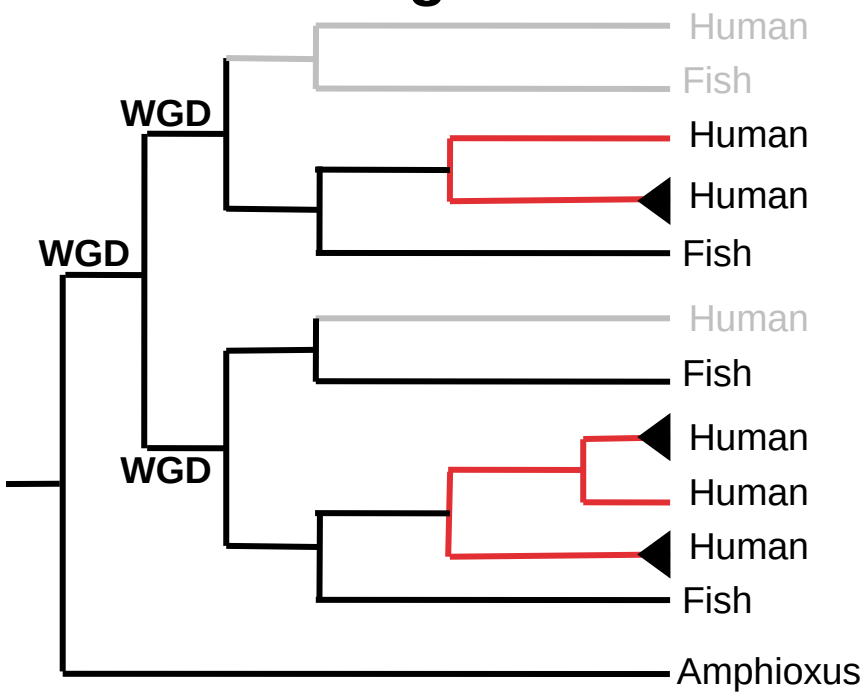
7,294 human ohnologs were identified.

High dosage sensitivity of ohnologs

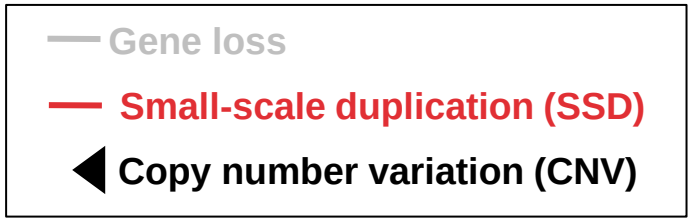
Ohnologs



Non-ohnologs



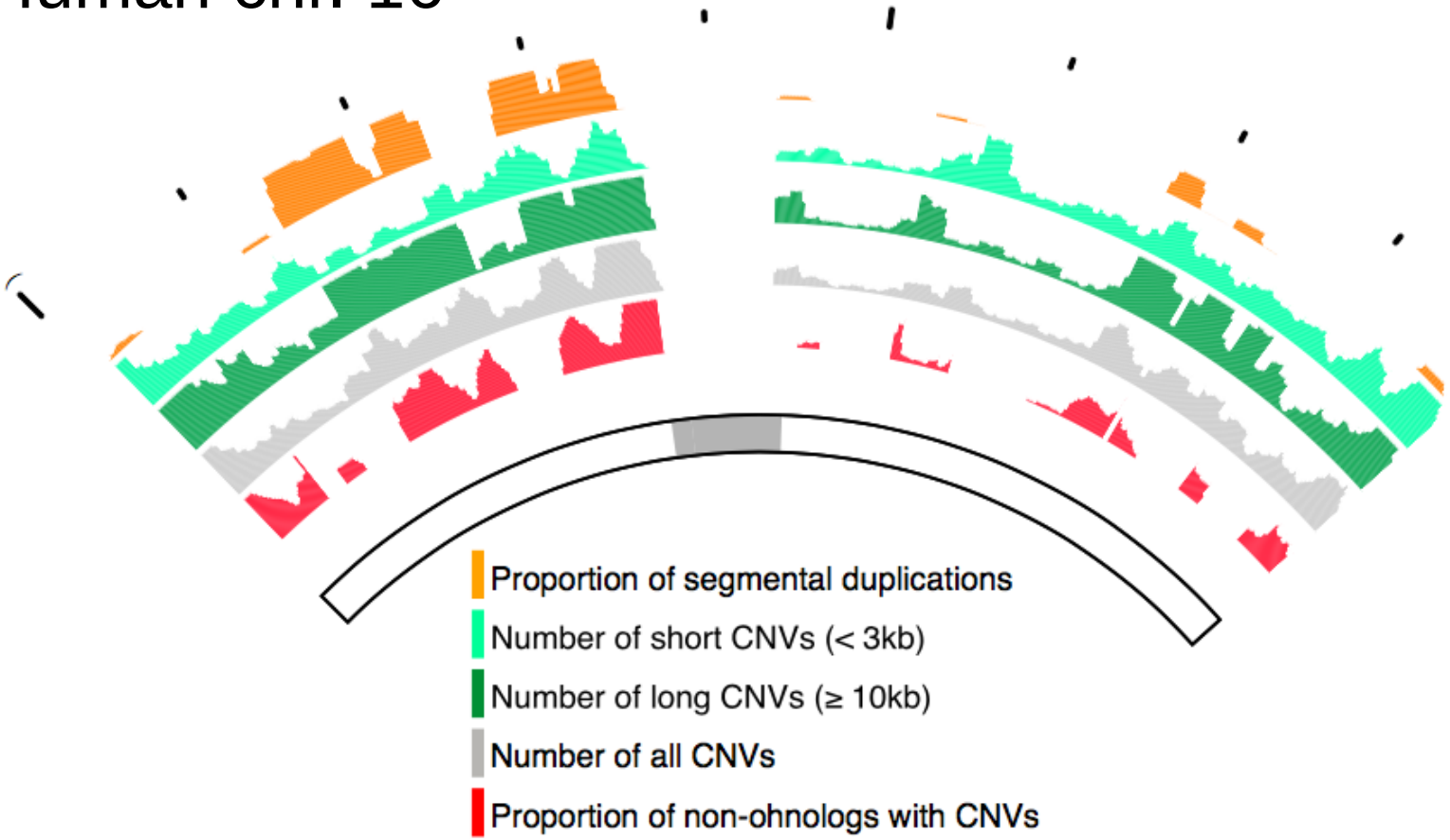
Makino and McLysaght. *PNAS* (2010)



Ohnologs have not experienced neither CNV nor SSD compared to non-ohnologs. ($P < 2.2 \times 10^{-16}$, χ^2 test)

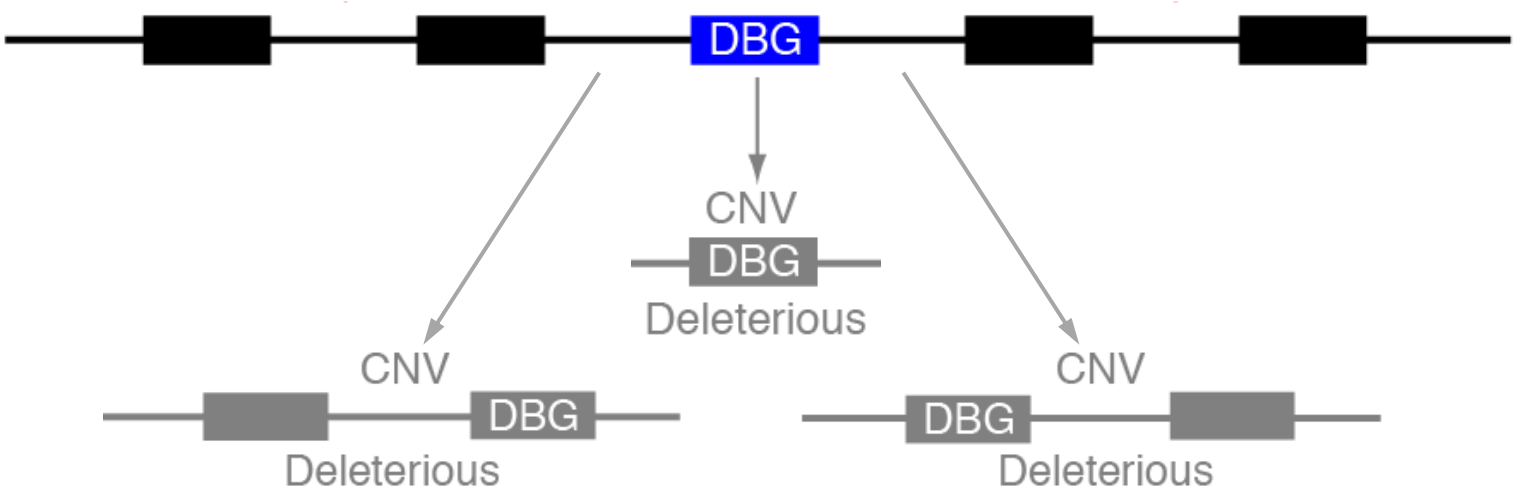
Biased distribution of CNVs

Human chr. 16



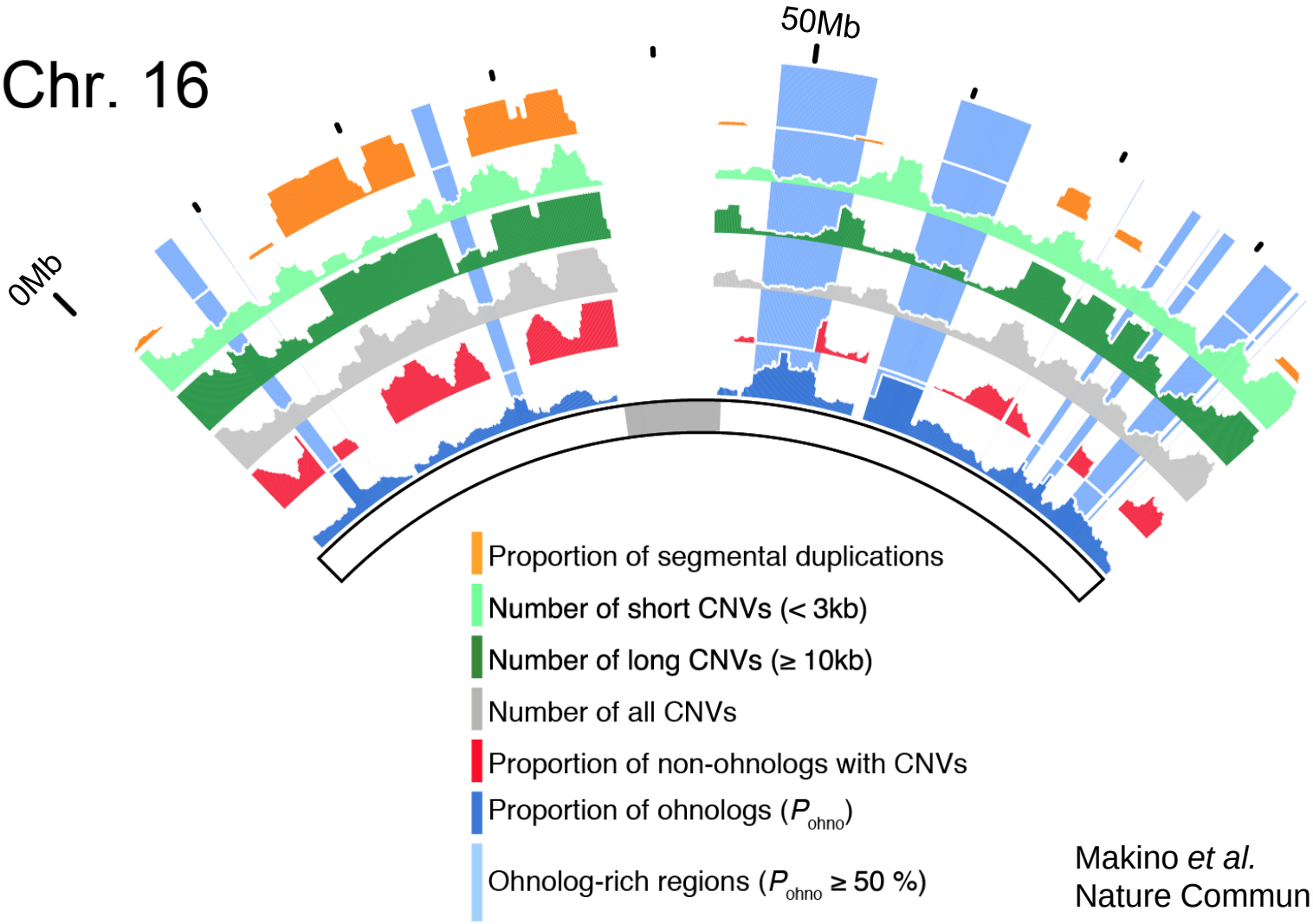
Makino *et al.*
Nature Communications (2013)

Deleterious CNVs of DBGs



Hypothesis: Genomic regions neighboring ohnologs are unlikely to experience CNVs.

CNV deserts in a human genome



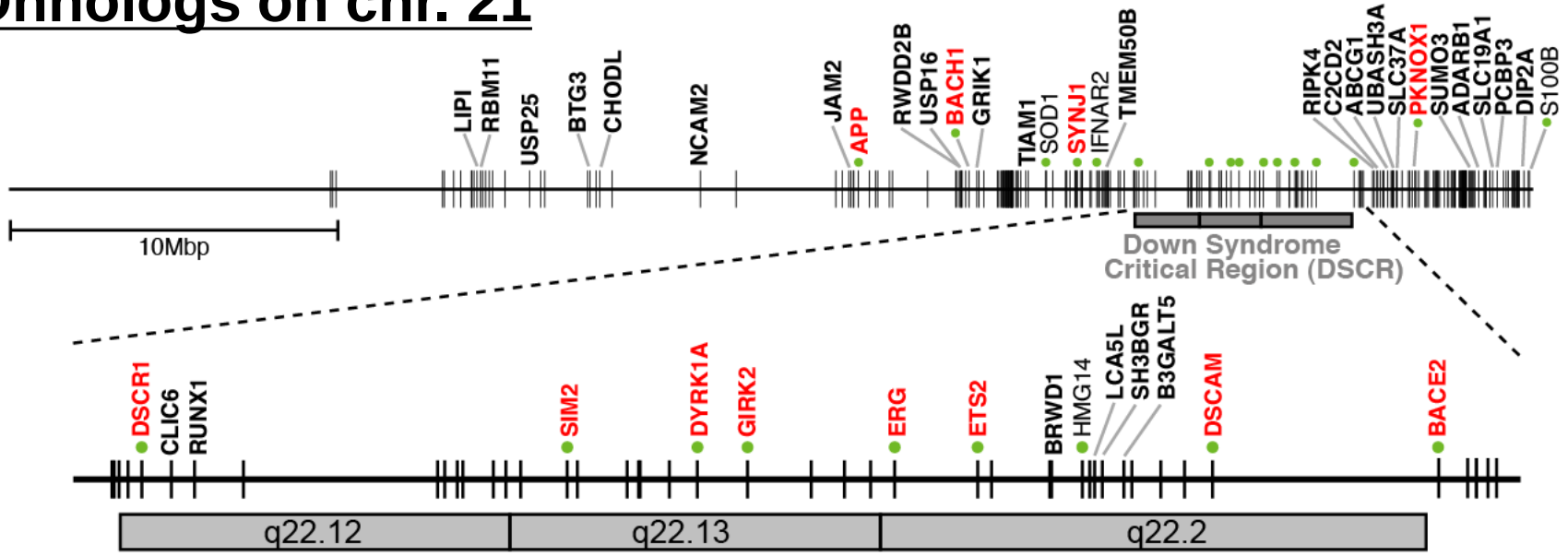
Makino *et al.*
Nature Communications (2013)

Ohnolog-rich regions are CNV deserts.

Trisomy 21: Down Syndrome

1.5-fold increase in dosage of all genes on chr. 21

Ohnologs on chr. 21



● Down Syndrome candidate genes

Makino and McLysaght. PNAS (2010)

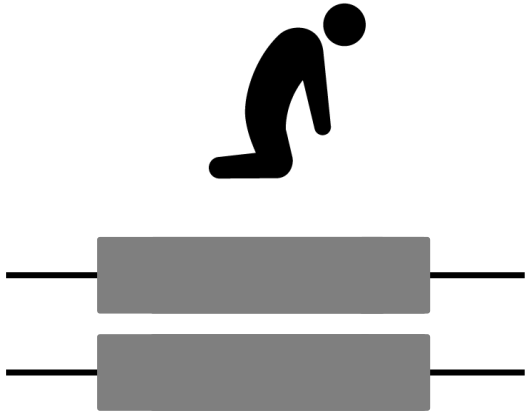
75% (12/16) of the candidate genes for this syndrome are ohnologs.

(P = 5.9 x 10^{-8}, fisher's exact test)

Pathogenic CNVs causing diseases

Healthy

Disease



Pathogenic CNV

http://pictogram2.com/?s=%E8%90%BD%E3%81%A1%E8%BE%BC%E3%82%80

Ohnologs overrepresented in pathogenic CNVs

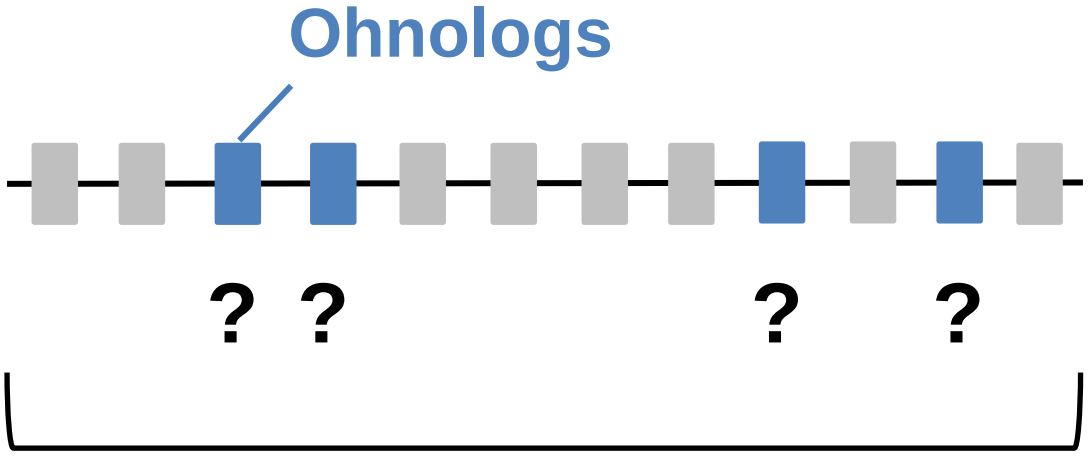
	Pathogenic CNVs			Control CNVs	
	Neuro-devl. diseases	Devl. delay (Cooper et al. Nature 2011)	Schizophrenia (Stefansson et al. Nature 2008)	WTCCC (>100kb)	HapMap (>100kb)
CNVs containing genes	15	51	22	66	194
CNVs containing ohnologs	14	47	20	20	69
Proportion	**0.93**	**0.92**	**0.91**	0.30	0.36

McLysaght, Makino *et al.* PNAS (2014)

> 90% of the pathogenic CNVs contained ohnologs.

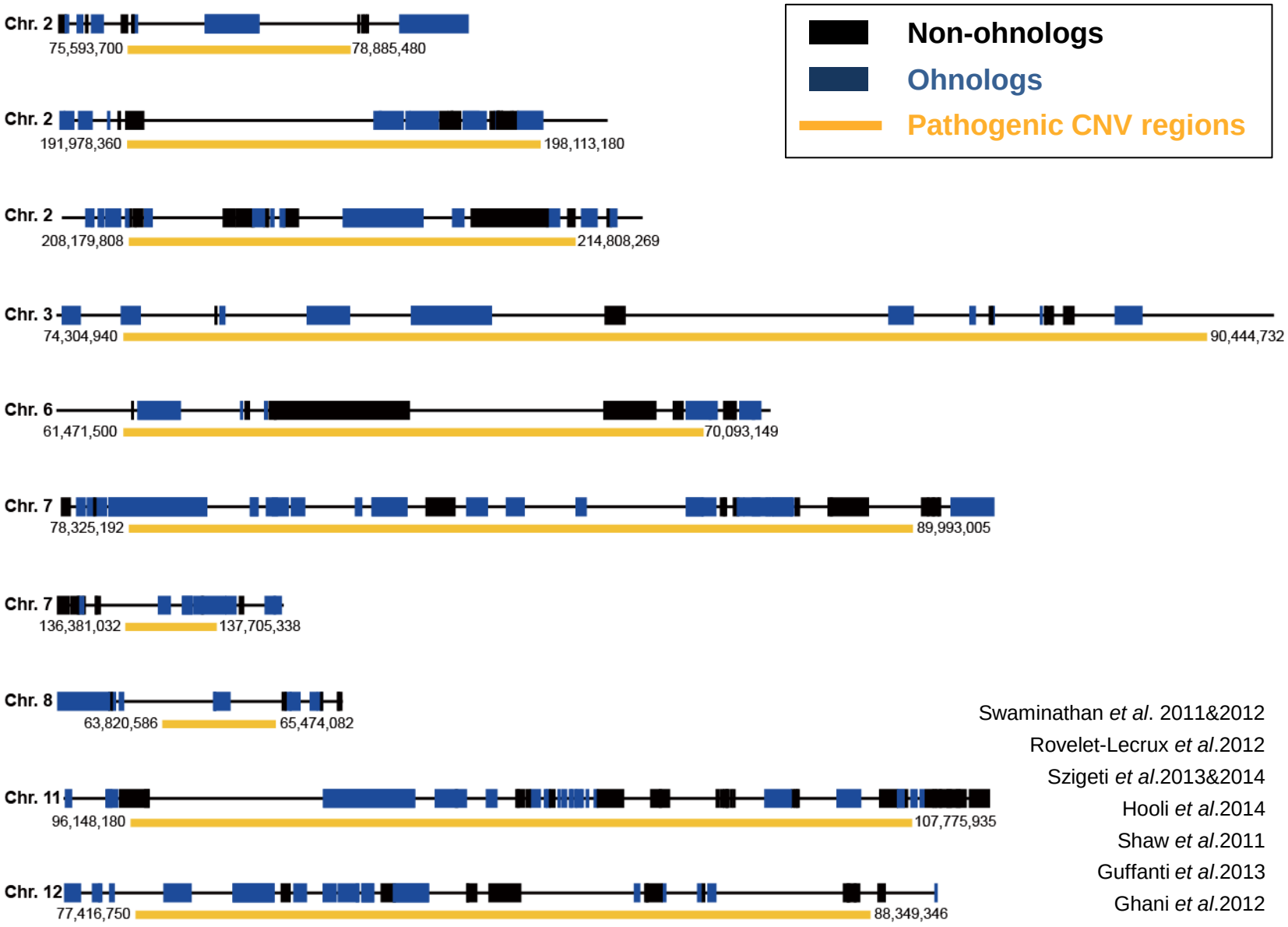
True causative gene on pathogenic CNVs

Which is a true disease causing gene in a pathogenic CNV?



Multiple ohnologs within
a pathogenic CNV

Genes on Alzheimer's Disease (AD) CNVs



Swaminathan *et al*. 2011&2012
Rovelet-Lecrux *et al*.2012
Szigeti *et al*.2013&2014
Hooli *et al*.2014
Shaw *et al*.2011
Guffanti *et al*.2013
Ghani *et al*.2012

Expression and phenotype for AD related genes

28 known Alzheimer's disease (AD) genes

Bertram *et al*.2008 Khanahmadi *et al*.2015
Giri *et al*.2016 Rosenburg *et al*.2016 AlzGene

64 ohnologs out of 248 genes in AD pathogenic CNVs

Swaminathan *et al*. 2011&2012 Shaw *et al*.2011
Rovelet-Lecrux *et al*.2012 Szigeti *et al*.2013&2014
Hooli *et al*.2014 Guffanti *et al*.2013 Ghani *et al*.2012

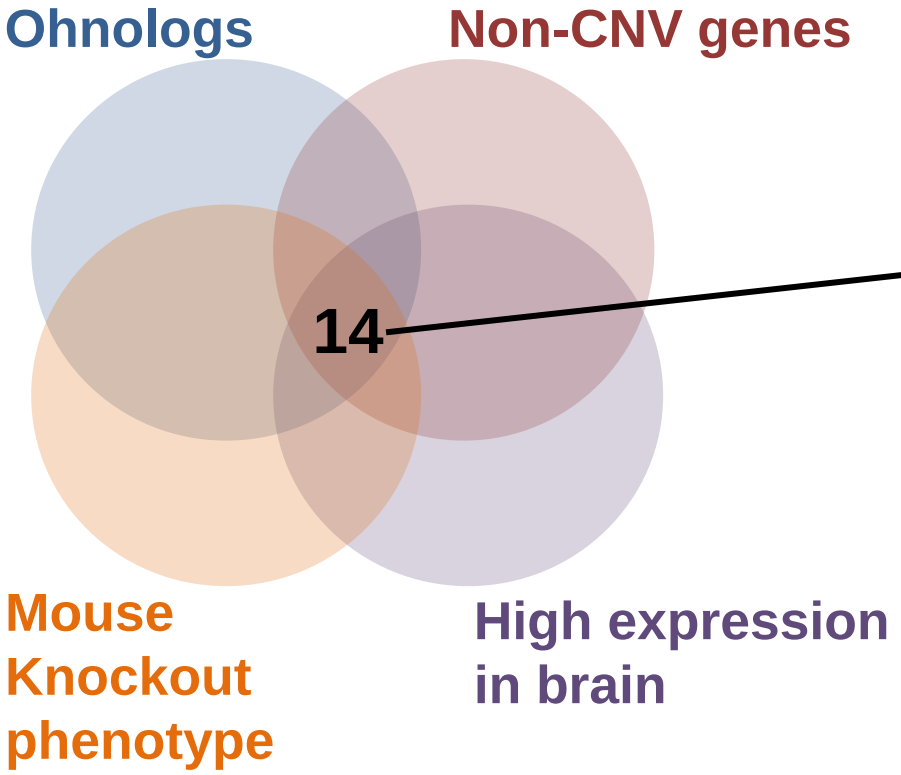
Expression pattern: Highly expressed in brain

Mouse knockout phenotype: Nervous system

Ohnologs in AD CNVs have high gene expression levels in brain and a nervous system related knockout phenotype.

Sekine and Makino. MBE,*in press*

Candidates of AD causing genes



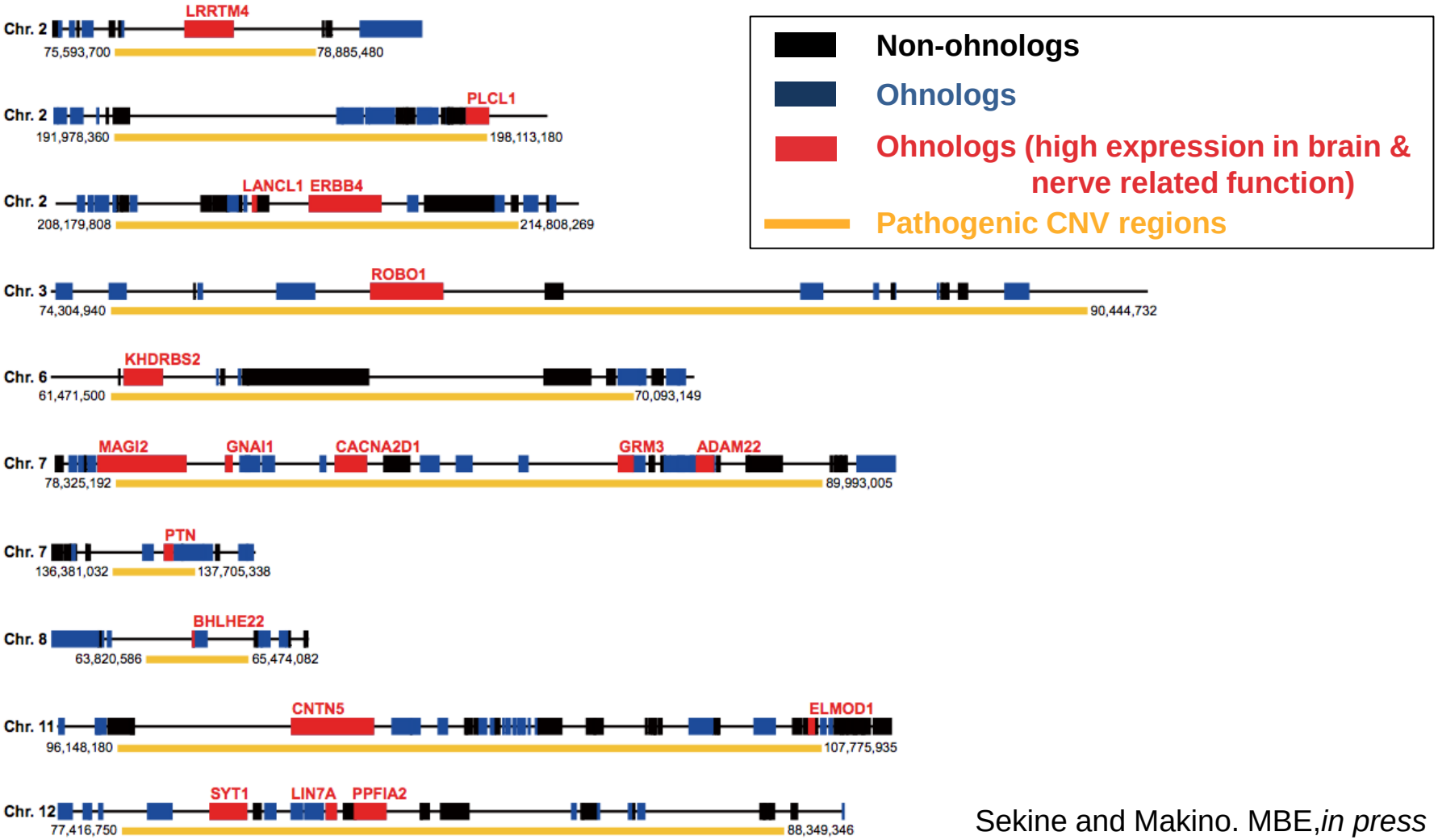
APP	***ATXN1***
BHLHE22	*CACNA2D1*
ERBB4	*GNAI1*
KHDRBS2	*LRRTM4*
MAGI2	*NCAM2*
NRXN1	*PTN*
ROBO1	*SLC30A3*

Sekine and Makino. MBE,*in press*

APP and ***ATXN1*** are previously reported AD genes.

Most of candidates are related to neuropsychiatric diseases.

Candidates of AD causing genes



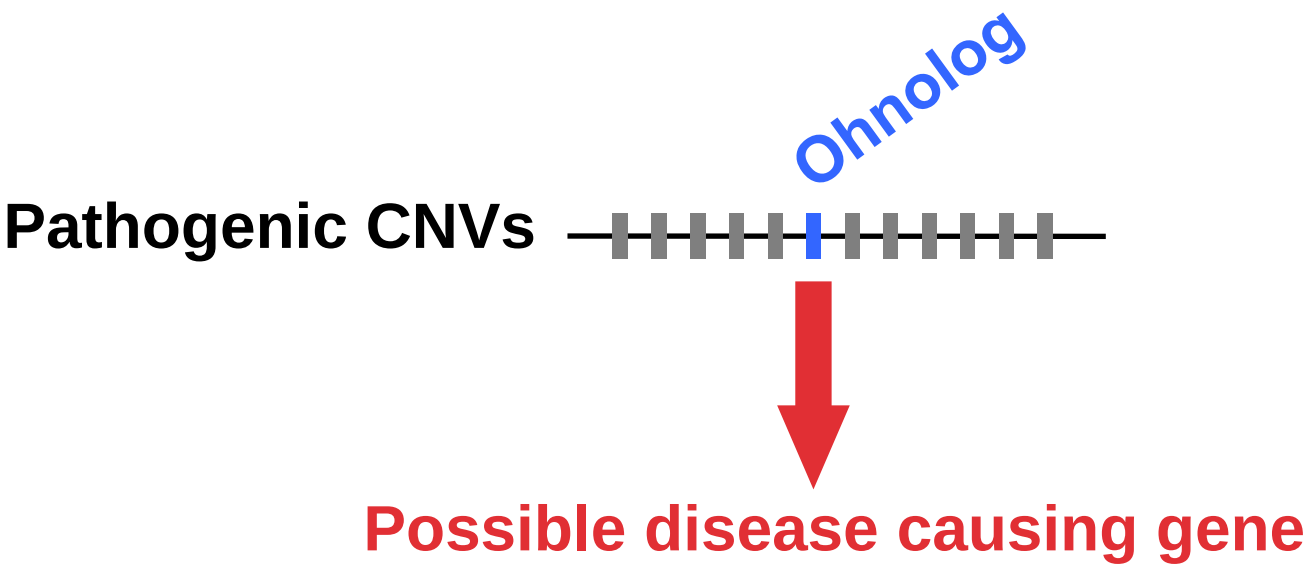
Sekine and Makino. MBE,*in press*

The estimation of causative genes is feasible even amongst multiple genes within AD CNV regions.

Conclusion

Ohnologs are dosage balanced and frequently associated with diseases.

Investigating CNV of ohnologs is an efficient means to find pathogenic CNV regions and disease causing genes within them.



Acknowledgements

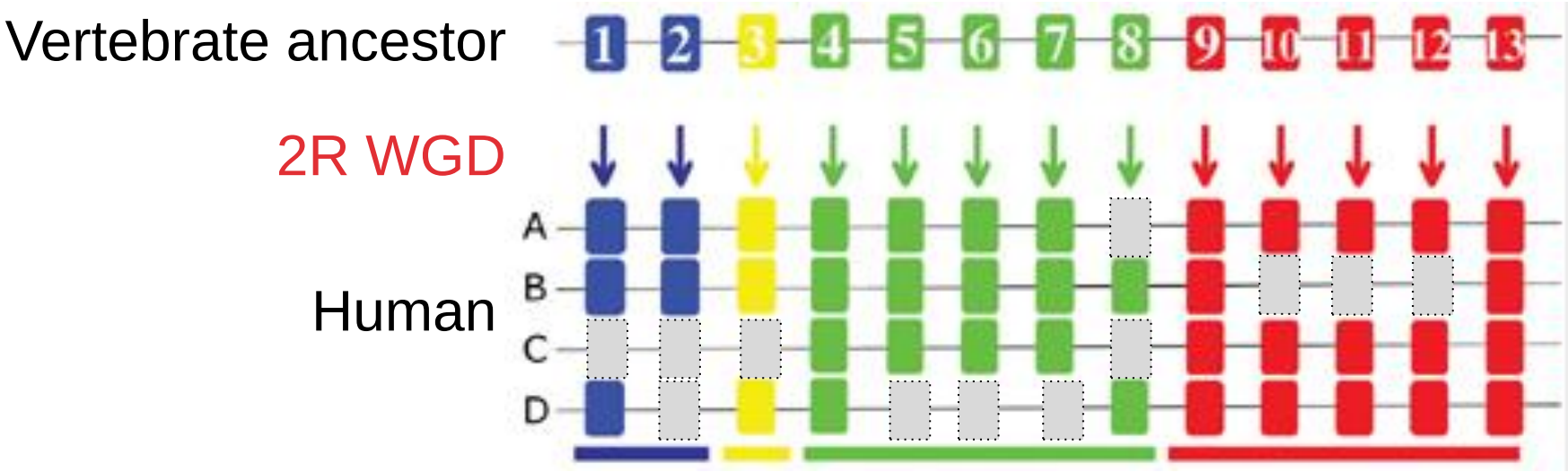
Aoife McLysaght (Trinity College Dublin, Ireland)

Sekine Mizuka (Tohoku University, Japan)

Duplicates generated by WGD

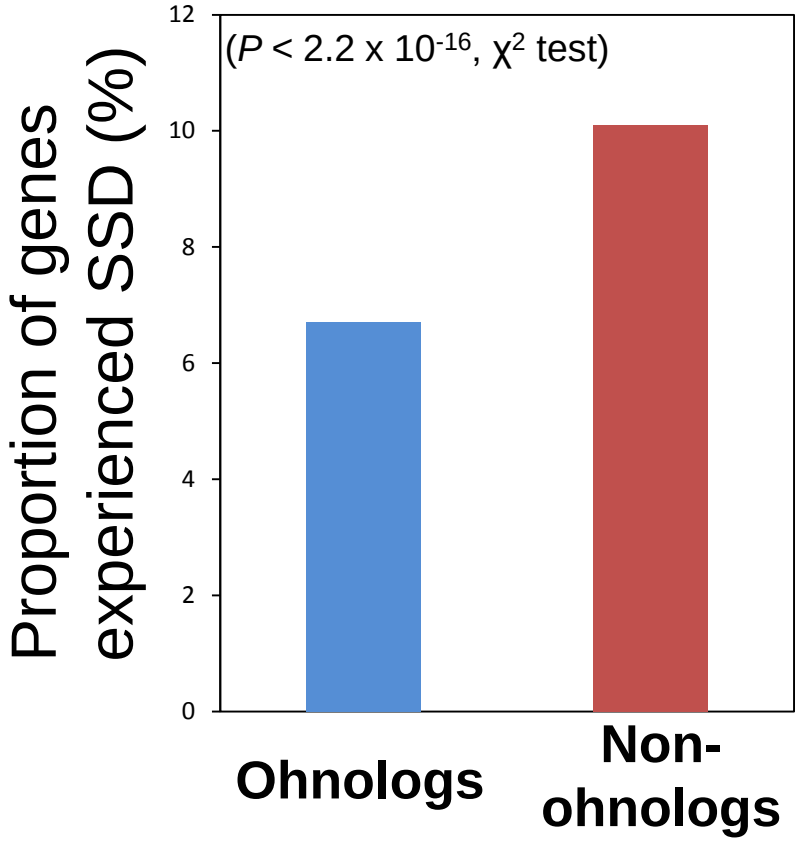
The number of genes increased fourfold by 2R WGD.

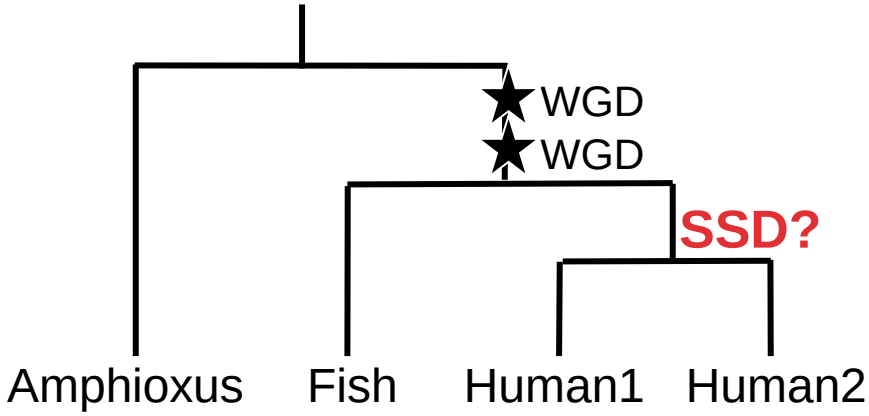
Typical example: Hox gene clusters



Garcia-Fernàndez Heredity (2005)

Less frequent SSD of ohnologs

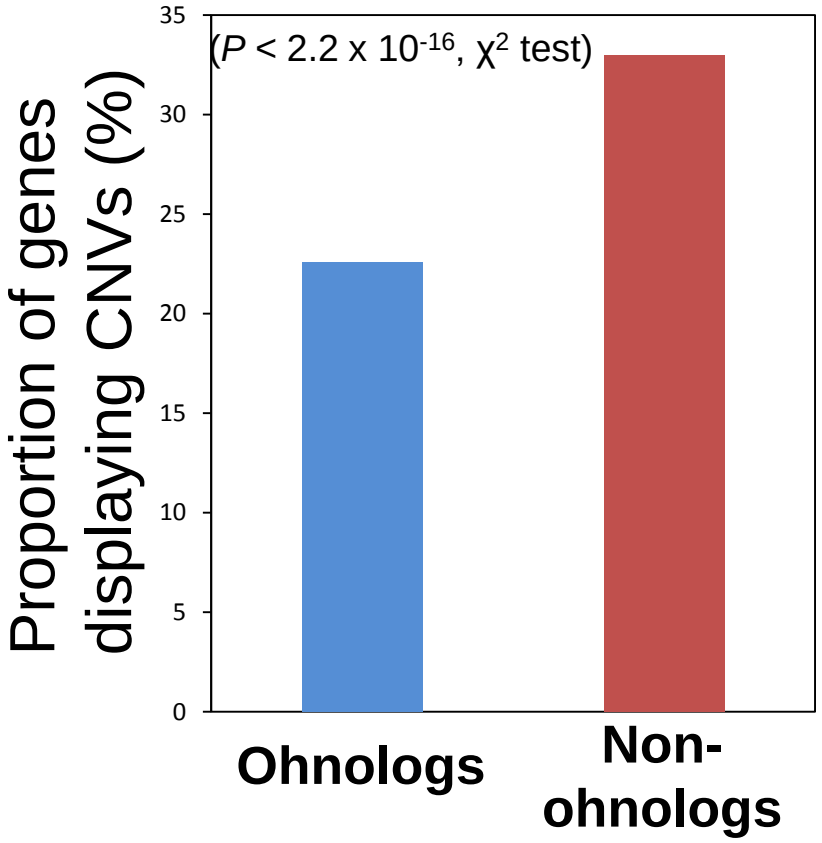




Ohnologs experienced SSD less frequently during evolution.

Makino and McLysaght. PNAS (2010)

Less frequent CNVs of ohnologs



CNVs in human populations

from Database of genomic variations (http://dgv.tcag.ca)

Ohnologs were unlikely to display CNVs in human populations.

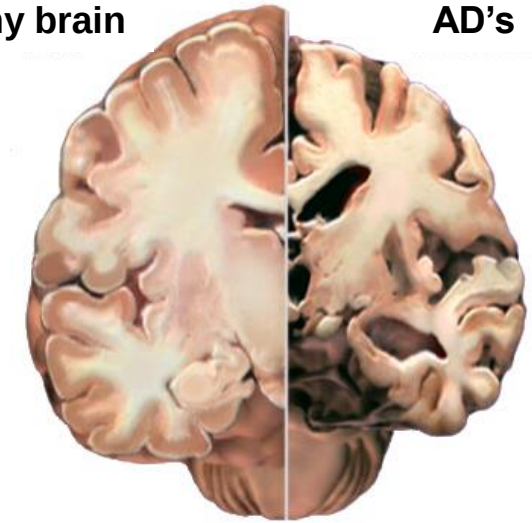
Makino and McLysaght. PNAS (2010)

Alzheimer's Disease (AD)

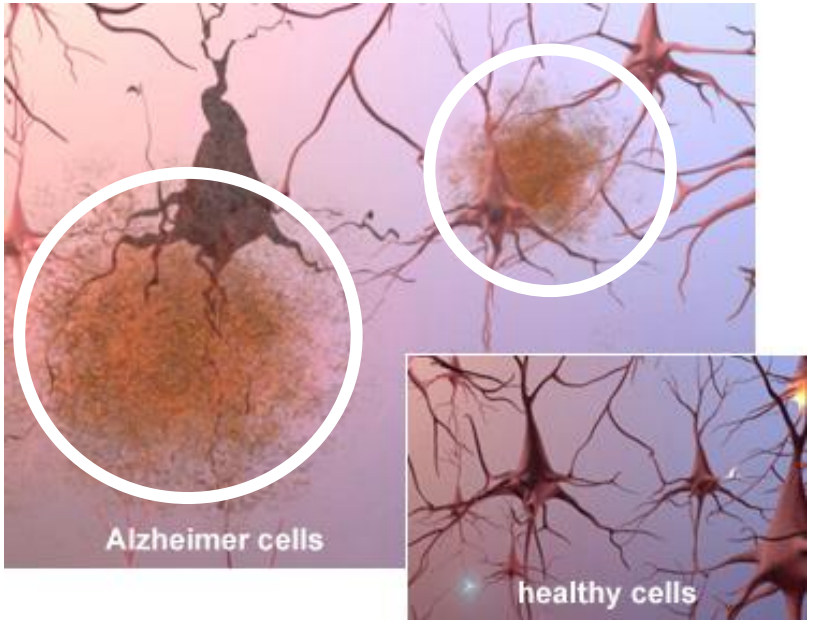
Caused by loss of neurons in the cortex and hippocampus impairment in memory, judgment, language, etc**.**

Mohammad *et al*. 2015

Healthy brain **AD's brain**



Accumulation of B-amyloid



alz.org

Alzheimer's Disease (AD) and control CNVs

64 AD causing CNVs (including 248 genes)

Shaw *et al*.2011

Swaminathan *et al*.2011

Ghani *et al*.2012

Rovelet-Lecrux *et al*.2012

Swaminathan *et al*.2012

Guffanti *et al*.2013

Szigeti *et al*.2013

Hooli *et al*.2014

Szigeti *et al*.2014

22,369 control CNVs

(CNVs containing genes)

Database of Genomic Variants

使用したデータ

AD CNVs 領域
61 CNVs / **497** 遺伝子
105 量感受性遺伝子
392 非量感受性遺伝子

Shaw *et al*.2011

Swaminathan *et al*.2011

Ghani *et al*.2012

Rovelet-Lecrux *et al*.2012

Swaminathan *et al*.2012

Swaminathan *et al*.2012

Guffanti *et al*.2013

Szigeti *et al*.2013

Hooli *et al*.2014

Szigeti *et al*.2014

既知のAD原因遺伝子
28 遺伝子

Bertram *et al*.2008

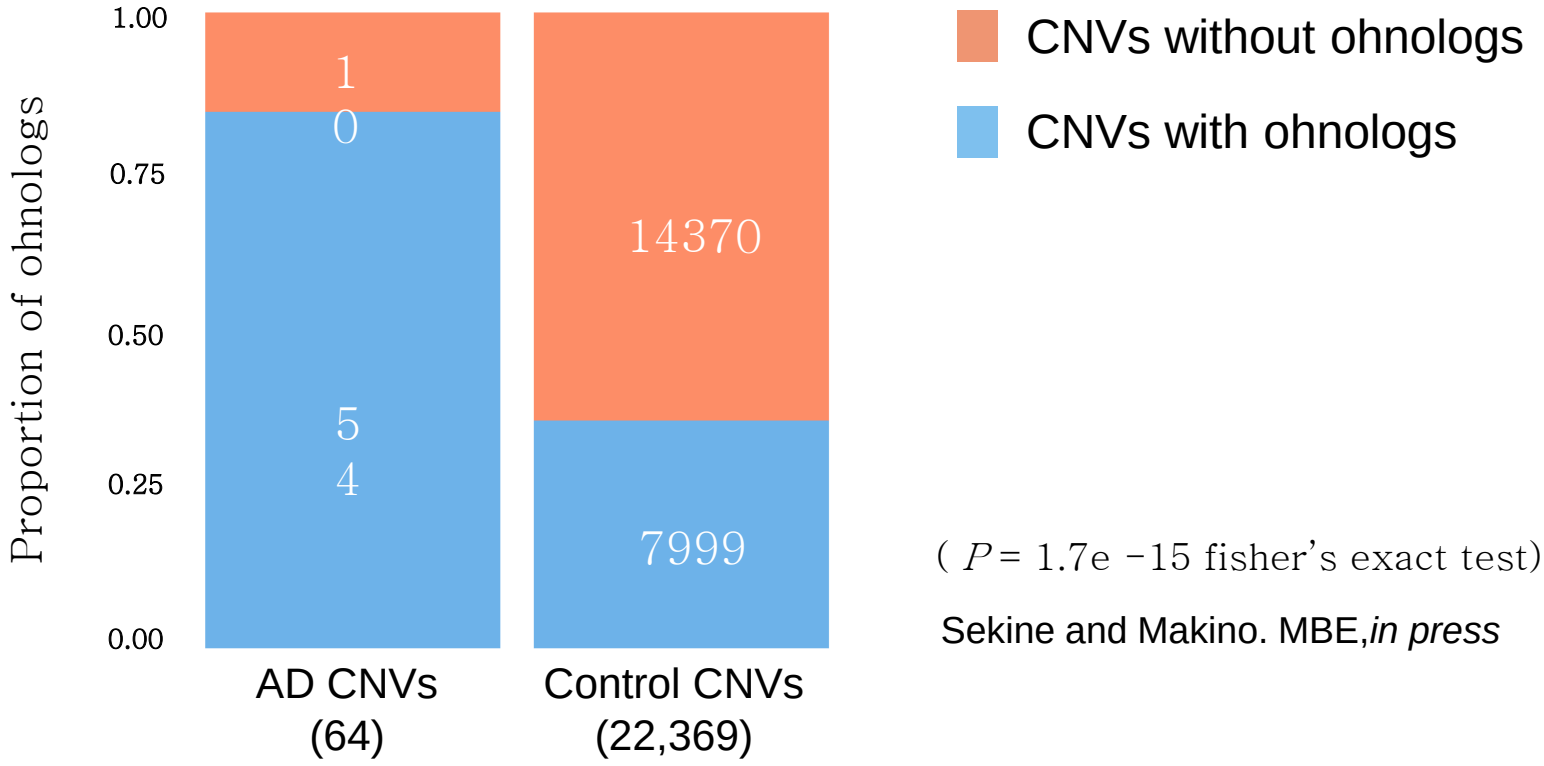
Khanahmadi *et al*.2015

Giri *et al*.2016

Rosenburg *et al*.2016

AlzGene

AD CNVs frequently including ohnologs

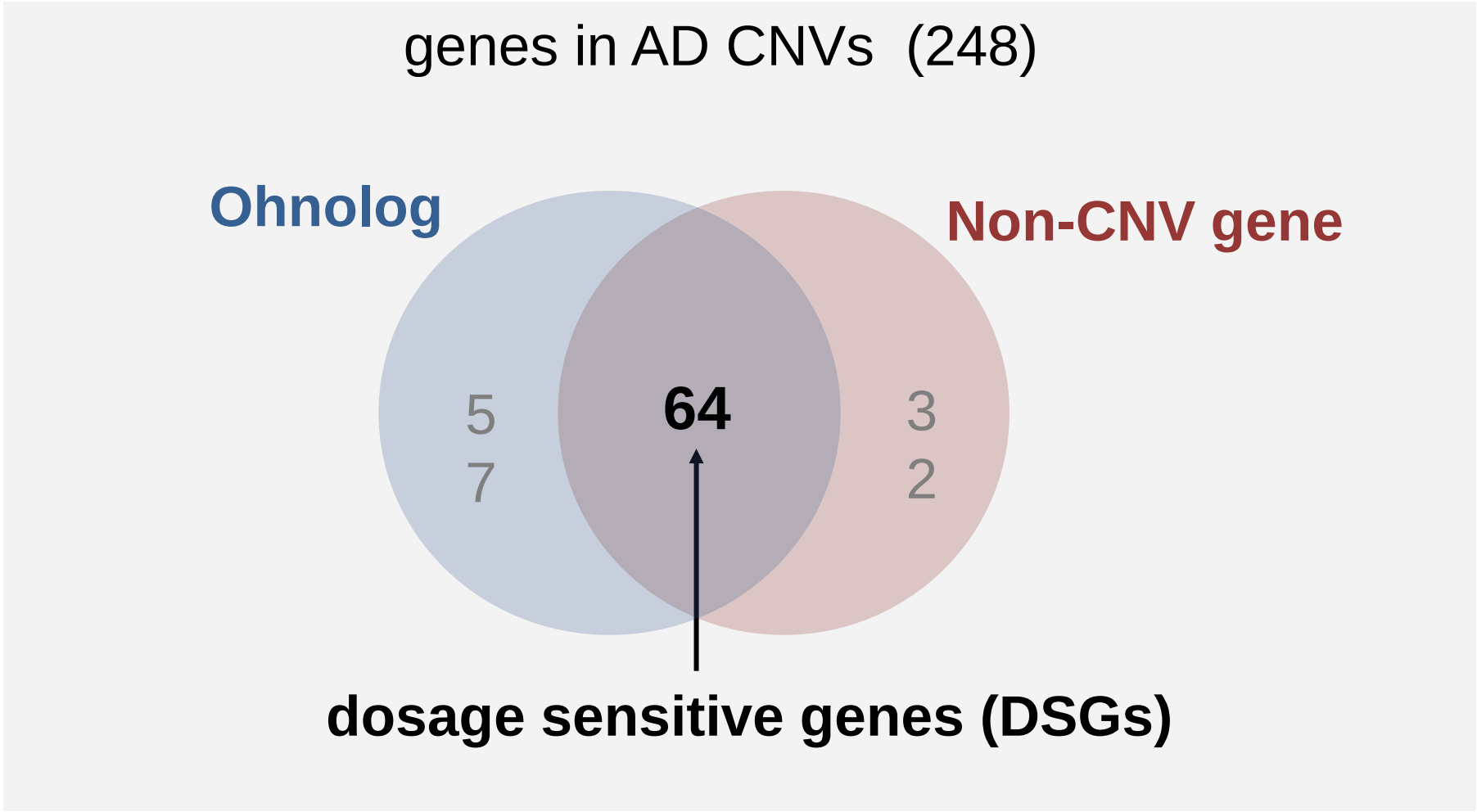


(P = 1.7e −15 fisher's exact test)

Sekine and Makino. MBE,*in press*

Proportion of CNVs having ohnologs for AD CNVs was higher than that for control CNVs.

Dosage sensitive genes on AD CNVs



Mouse knockout phenotype

Known AD genes (28)

Rosenburg *et al*.2016 Bertram *et al*.2008 Shen *et al*.2016
OMIM Giri *et al*.2016 Khanahmadi *et al*.2015 AlzGene

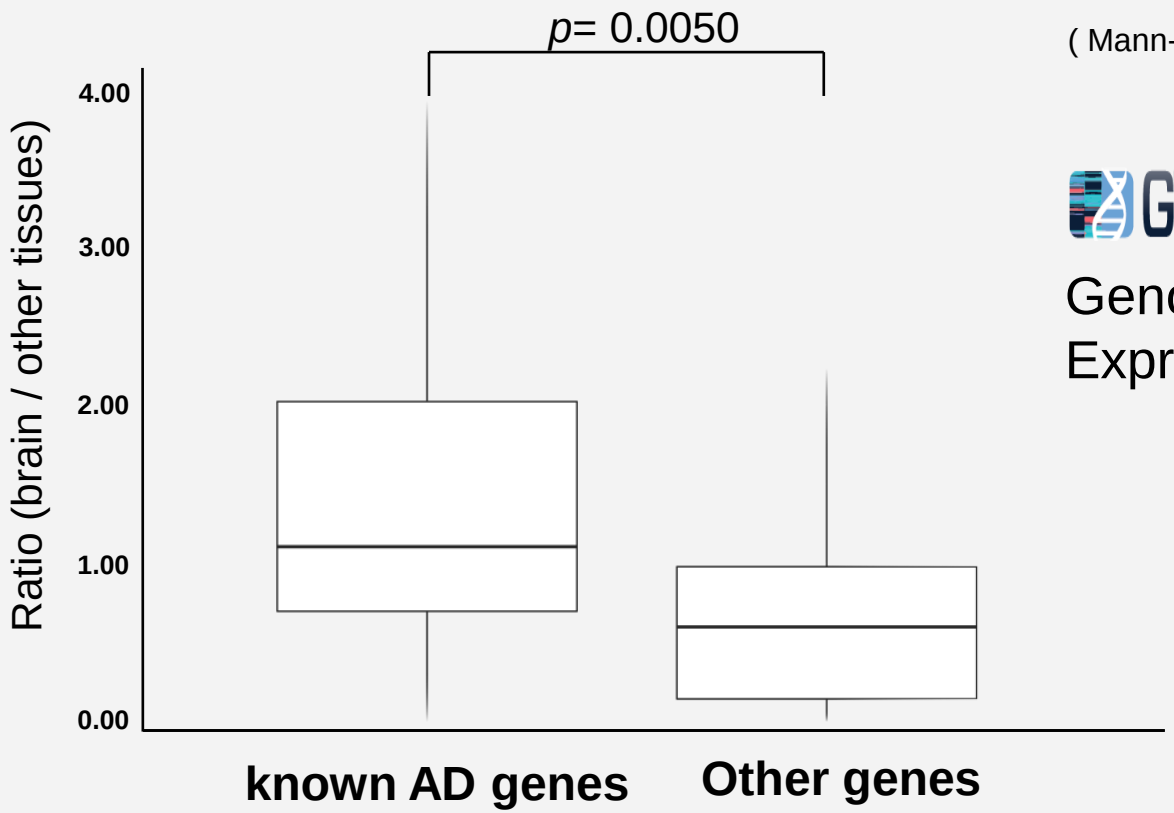
Phenotype term	*P* value
nervous system	0.014
behavior / neurological	0.13
growth /size/ body	0.57
muscle	0.68
….	

Dosage sensitive genes in AD CNVs (64)

Phenotype term	*P* value
nervous system	0.013
….	

Expression of known AD genes

Known AD genes had high expression in brain.



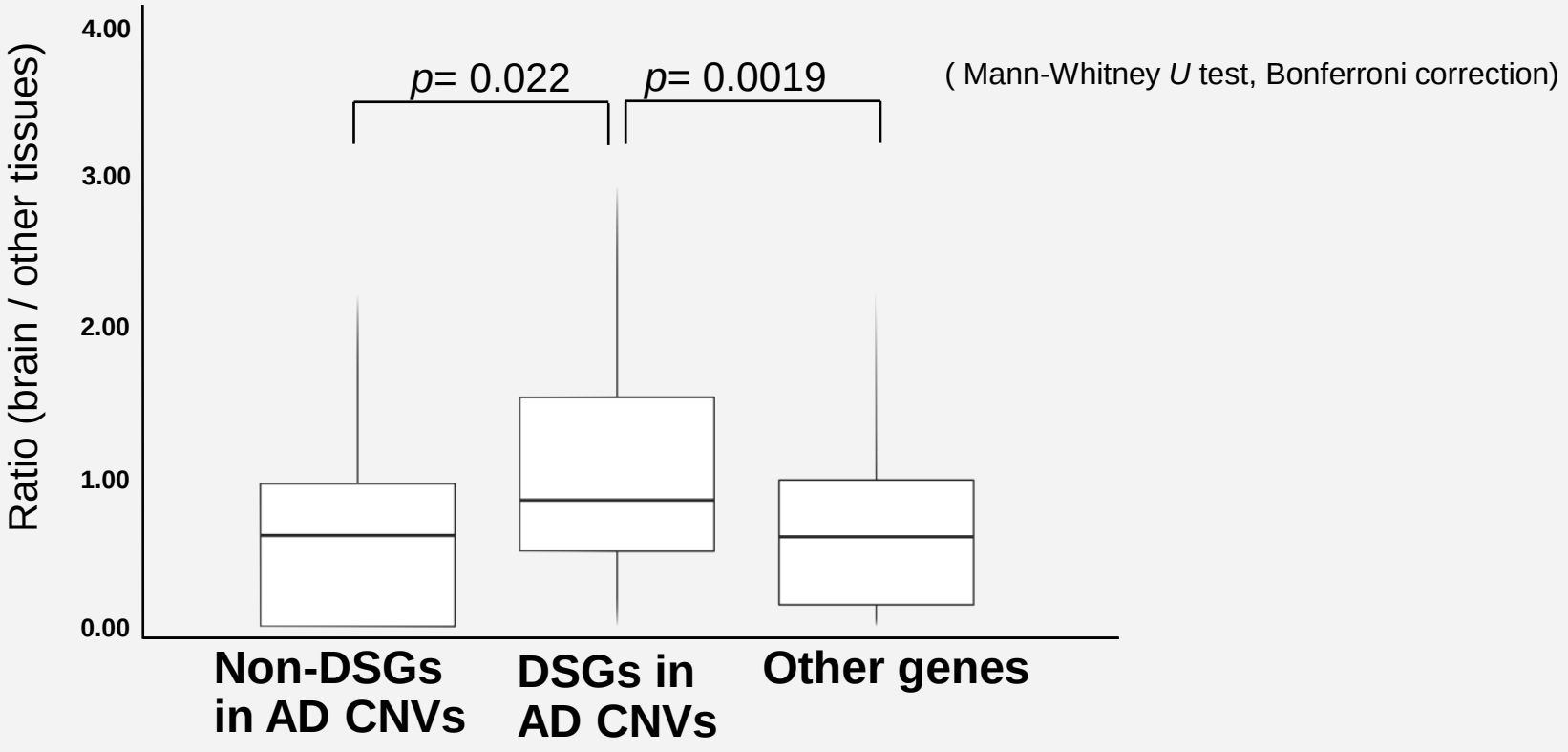
(Mann-Whitney *U* test, Bonferroni correction)

GTEx Portal

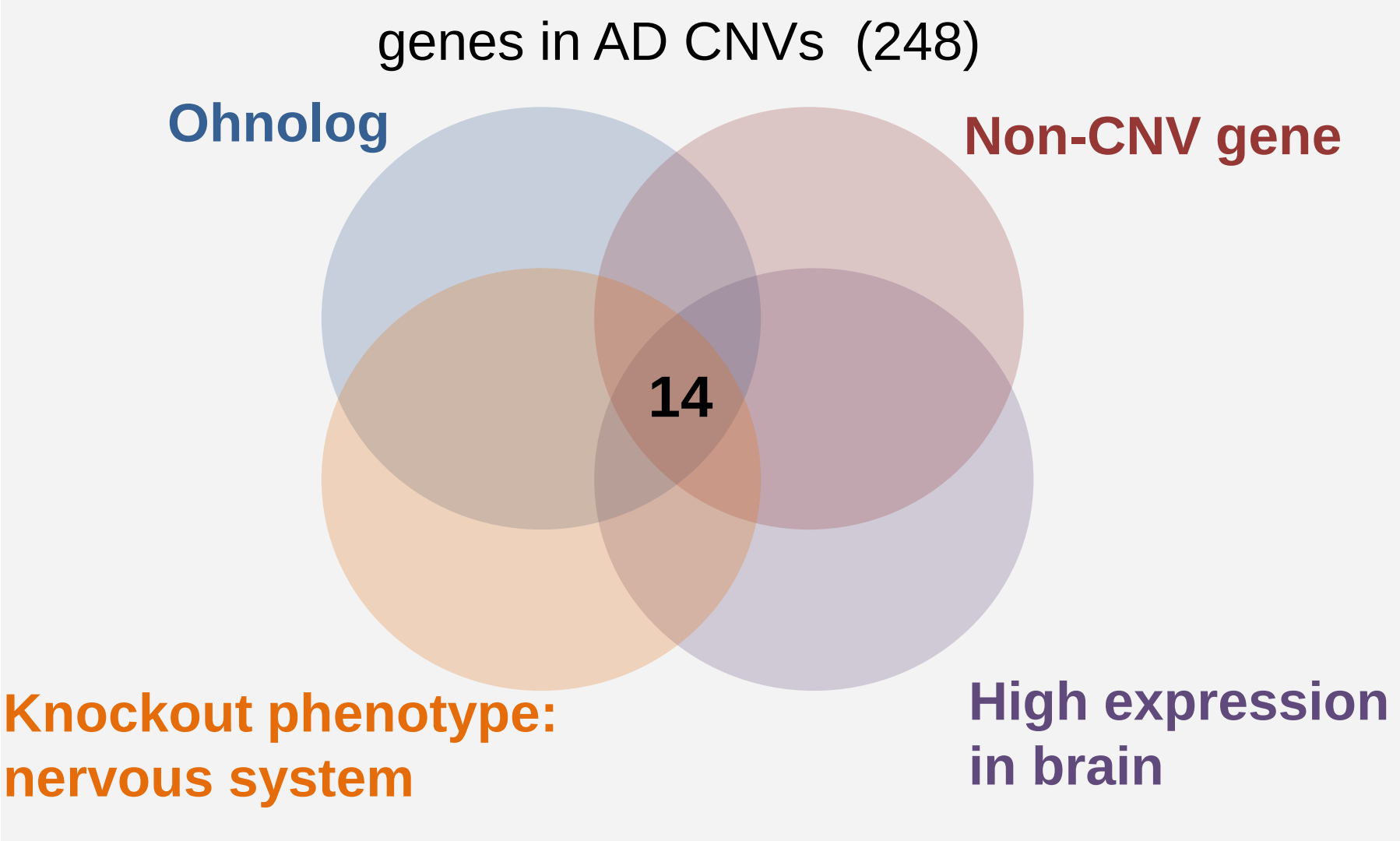
Genotype-Tissue Expression portal

Expression of DSGs in AD CNVs

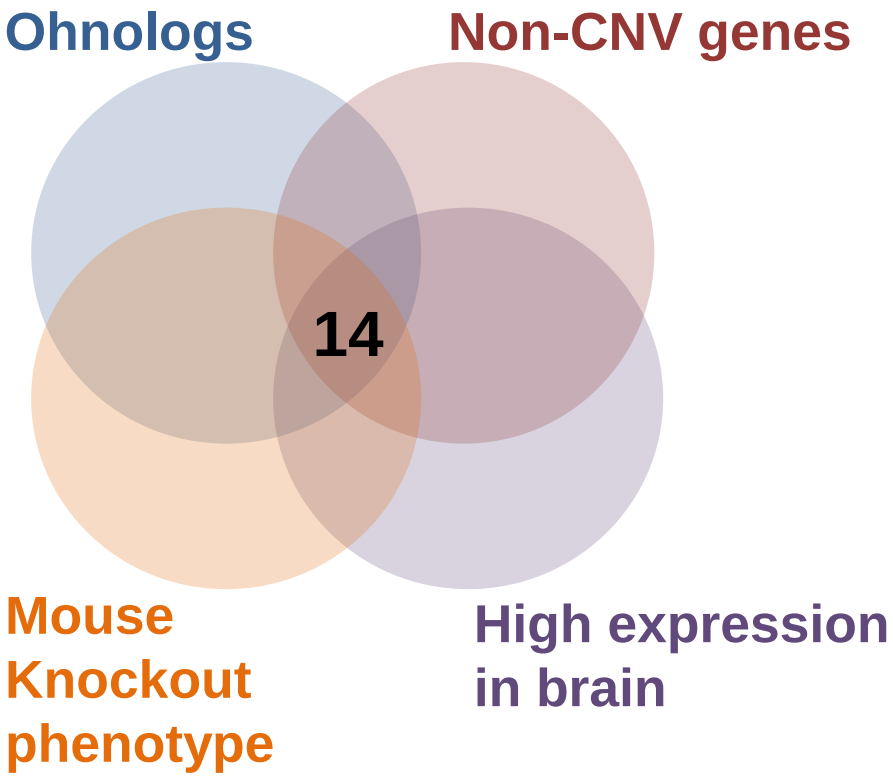
Dosage sensitive genes (DSGs) in AD CNVs had high expression in brain.



Candidates of AD causing genes



Candidates of AD causing genes



APP	***ATXN1***
BHLHE22	*CACNA2D1*
ERBB4	*GNAI1*
KHDRBS2	*LRRTM4*
MAGI2	*NCAM2*
NRXN1	*PTN*
ROBO1	*SLC30A3*

Sekine and Makino. MBE,*in press*

APP and ***ATXN1*** are previously reported AD genes.

Most of candidates are related to neuropsychiatric diseases.