

Genome-wide association study provides insights into the etiology and epidemiology of intracranial germ cell tumors.

Lead Presenter:

Department of Genome Informatics,
Graduate School of Medicine, the University of Tokyo

Kyuto Sonehara (曾根原 究人)

Co-presenter: Yui Kimura, Yoshiko Nakano, Tatsuya Ozawa, Meiko Takahashi, Ken Suzuki, Takashi Fujii, Yuko Matsushita, Arata Tomiyama, Toshihiro Kishikawa, Kenichi Yamamoto, Tatsuhiko Naito, Tomonari Suzuki, Shigeru Yamaguchi, Tomoru Miwa, Hikaru Sasaki, Masashi Kitagawa, Naoyuki Ohe, Junya Fukai, Hideki Ogiwara, Atsufumi Kawamura, Satoru Miyawaki, Fumihiko Matsuda, Nobutaka Kiyokawa, Koichi Ichimura, Ryo Nishikawa, Yukinori Okada, Keita Terashima

COI Disclosure Information

Lead Presenter: Kyuto Sonehara

Principal Researcher: Ryo Nishikawa, Yukinori Okada, Keita Terashima

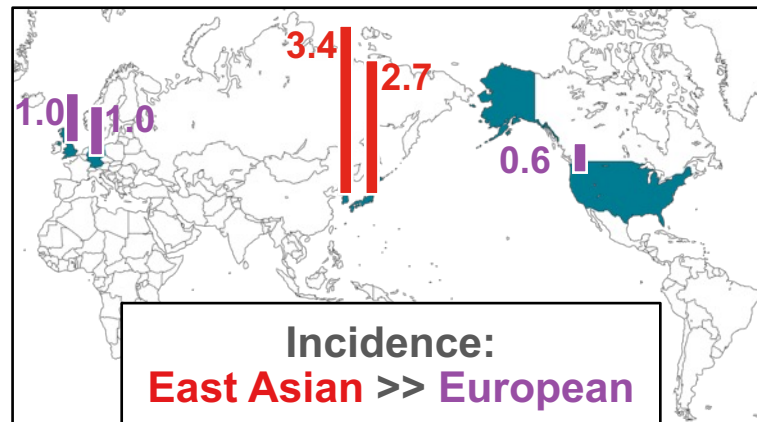
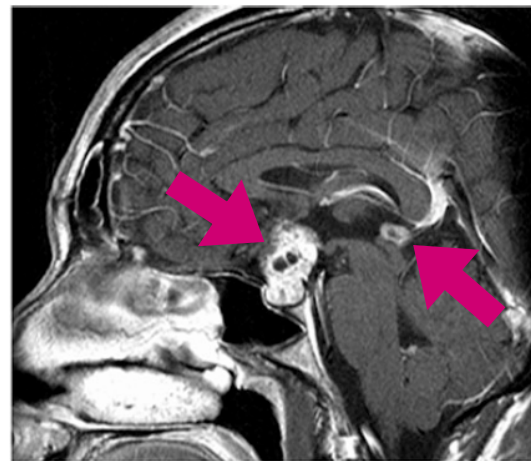
We have no financial relationships to disclose.

Intracranial Germ Cell Tumors: IGCTs

- The **second most common** (~12%) pediatric brain tumor in Japan.
- **High incidence in East Asia.**
(e.g., Japan, 2.7/million/year
U.S., 0.6/million/year)
- **The low incidence** has limited basic research on this rare disease.

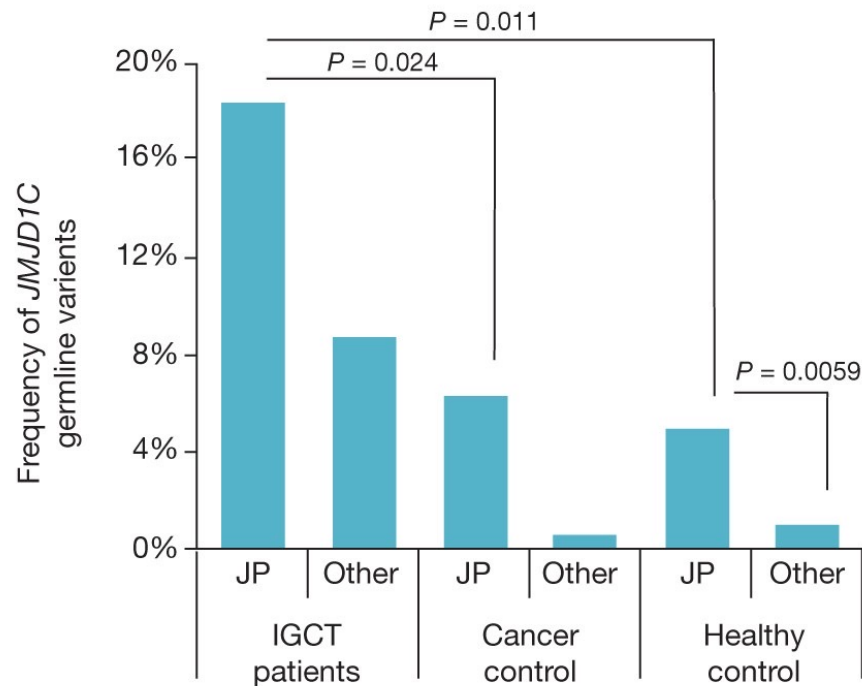


Elucidation of IGCTs etiology is an important challenge especially **for East Asia including Japan.**



Intracranial Germ Cell Tumors: IGCTs

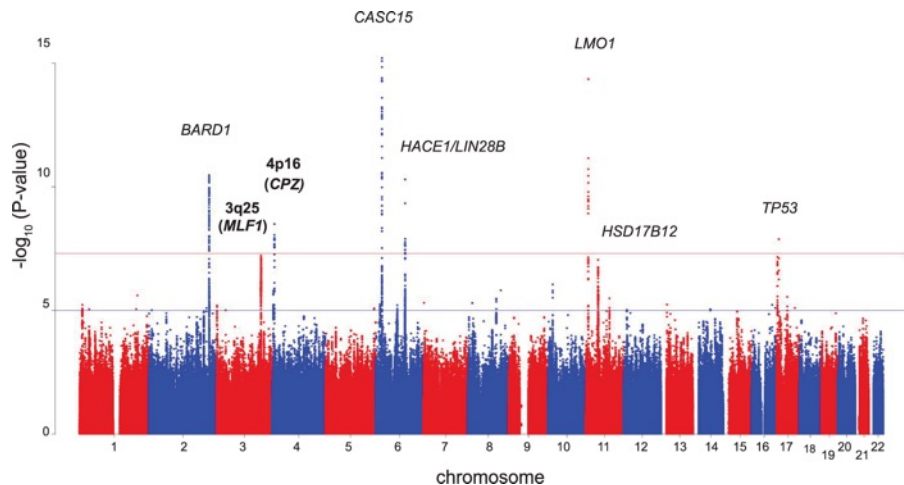
- The largest earlier study on IGCTs germline genetics involved whole-exome sequencing on 62 cases.
- Rare coding variants of *JMJD1C* were implicated in the susceptibility.
- Genome-wide screening of susceptibility genes has never been conducted.



Common genetic variants and pediatric tumors

- Genome-wide association studies reveal that pediatric tumors without apparent familiarity also have genetic predisposition.

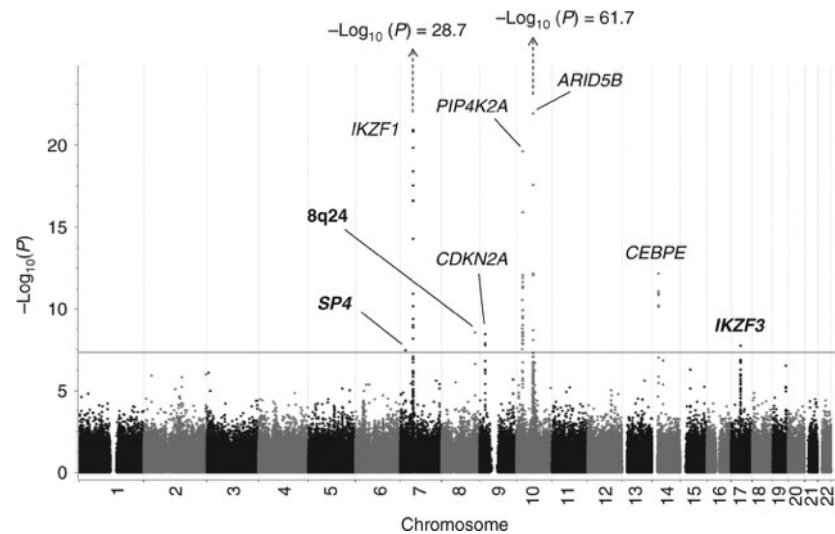
Neuroblastoma



Diskin SJ et al. *Nat Genet* 2012

McDaniel LD et al. *PLoS Genet* 2017

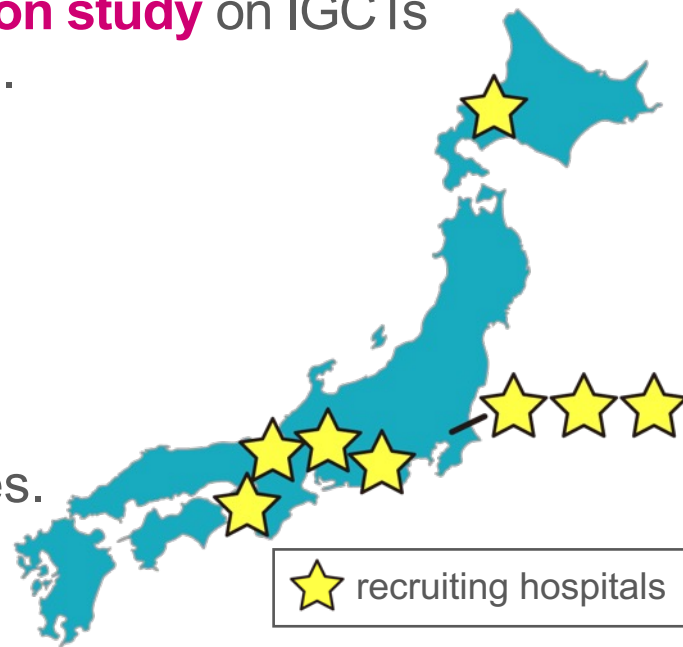
Childhood acute lymphoblastic leukemia



Wiemels JL et al. *Nat Commun* 2018

Study Design

- We designed **an initial genome-wide association study** on IGCTs with **the largest case sample size** ever reported.
- **138 IGCTs cases**: from the National Center for Child Health and Development and seven other hospitals throughout Japan.
- **808 Healthy controls**: from the University of Tokyo, Osaka University, and affiliated institutes.
- In addition, **99 IGCTs cases** and **1,026 controls** were available as a replication dataset.



(from Ichimura K et al. *Acta Neuropathol* 2016 and Okada Y et al. *Nat Commun* 2018)

Study Design

Discovery GWAS

 138 cases  808 controls

Quality control filter

 133 cases  762 controls

497,059 high-quality genotyped SNPs

Whole-genome
genotype imputation

8,530,563 genetic variants for GWAS

Functional characterization by *in silico* and *in vitro* analyses

Replication Analysis

 99 cases  1,026 controls

Replication analysis with independent data

Significant loci
($P < 5.0 \times 10^{-8}$)

Study Design

Discovery GWAS

 138 cases  808 controls

Quality control filter

 133 cases  762 controls

497,059 high-quality genotyped SNPs

Whole-genome
genotype imputation

8,530,563 genetic variants for GWAS

Functional characterization by *in silico* and *in vitro* analyses

Replication Analysis

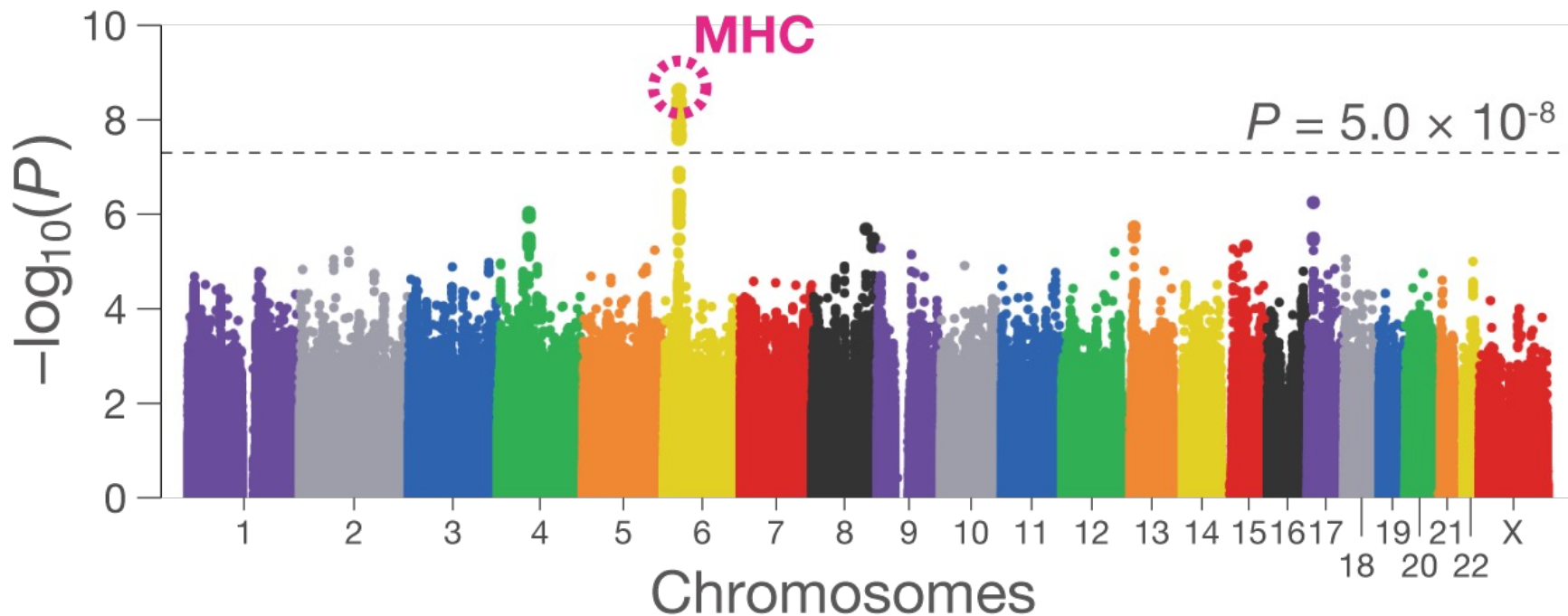
 99 cases  1,026 controls

Replication analysis with independent data

Significant loci
($P < 5.0 \times 10^{-8}$)

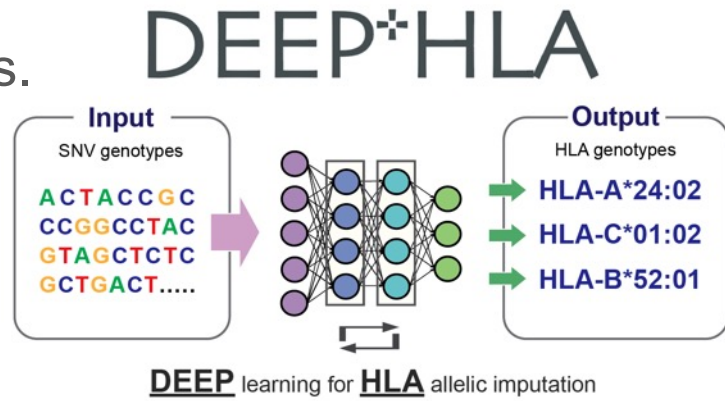
Genome-Wide Association Study

Genome-wide significant variants were found at the **major histocompatibility complex region (MHC; $P = 2.4 \times 10^{-9}$)**.

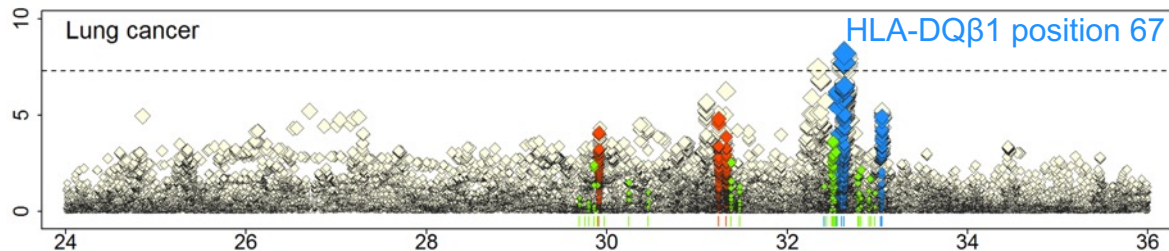


HLA imputation analysis

- Human leukocyte antigen (HLA) variants in MHC have been implicated in various cancers.
- Typically, MHC associations are fine-mapped to HLA variants (e.g., HLA-DQ β 1 with lung cancer).
- We computationally inferred HLA variant genotypes based on the surrounding SNPs using a convolutional deep learning method.



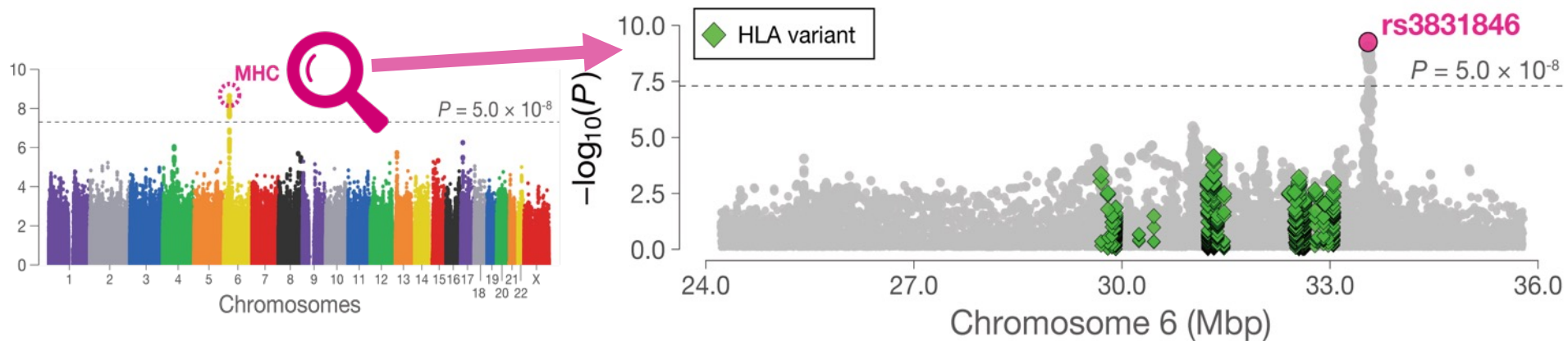
Naito T et al. *Nat Commun* 2021



Hirata J et al. *Nat Genet* 2019

No association with HLA variants

- Performing HLA imputation, we analyzed HLA variants association.
- Unexpectedly, **no HLA variants showed association**.
- The lead variant was **rs3831846**, a **4-bp deletion**.



Study Design

Discovery GWAS

 138 cases  808 controls

Quality control filter

 133 cases  762 controls

497,059 high-quality genotyped SNPs

Whole-genome
genotype imputation

8,530,563 genetic variants for GWAS

Functional characterization by *in silico* and *in vitro* analyses

Replication Analysis

 99 cases  1,026 controls

Replication analysis with independent data

Significant loci
($P < 5.0 \times 10^{-8}$)

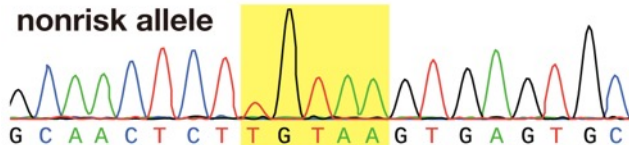
Validating Deletion and Replicating Association

Discovery Dataset
(133 cases vs 762 controls)

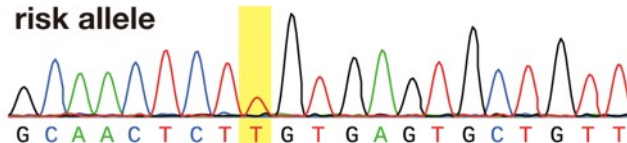
SNP	Chr	Position	Alleles	Risk Allele	Freq. in Cases	Freq. in Controls	Odds Ratio (95% CI)	P-value
rs3831846	6	33,548,346	<u>TGTA</u> A/T	T	0.63	0.43	2.46 (1.83–3.31)	2.4×10^{-9}

Sanger sequencing validation

nonrisk allele



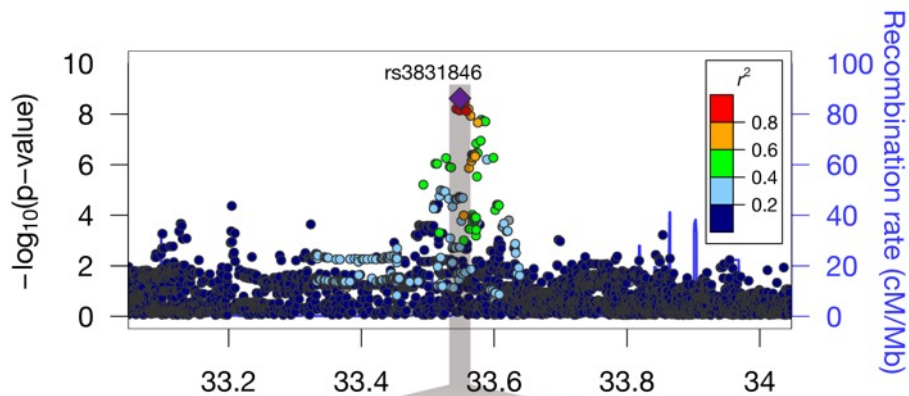
risk allele



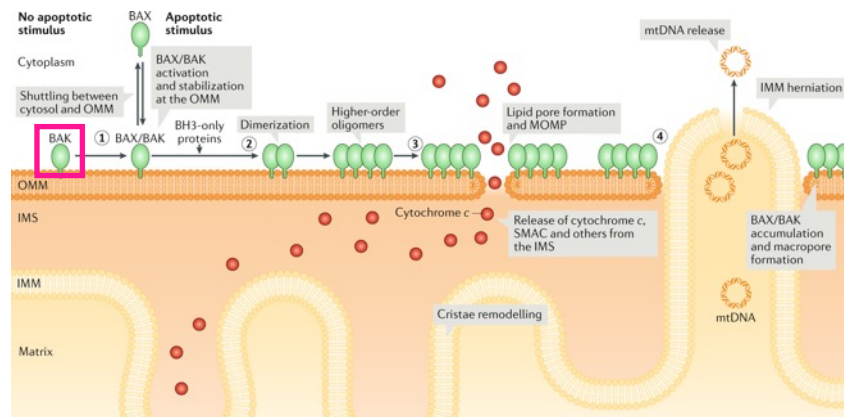
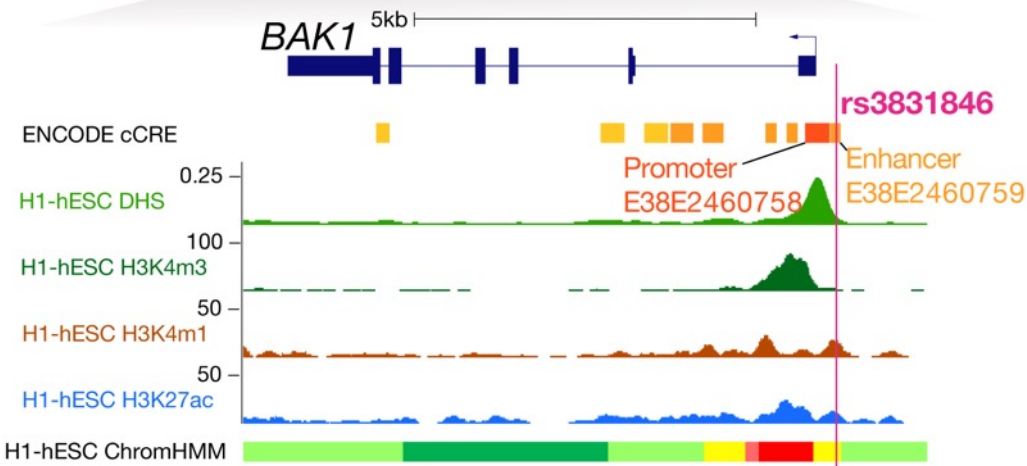
Replication Dataset
(99 cases vs 1,026 controls)

Freq. in Cases	Freq. in Controls	Odds Ratio (95% CI)	P-value
0.62	0.42	2.22 (1.63–3.03)	1.7×10^{-7}

rs3831846 is in *BAK1* enhancer



- Rs3831846 resides in an **enhancer** element adjacent to the ***BAK1*** promoter.
- ***BAK1*** (BCL2 Antagonist/Killer 1): an **apoptosis-inducer** of the BCL2 protein family, the apoptosis regulators.



Study Design

Discovery GWAS

 138 cases  808 controls

Quality control filter

 133 cases  762 controls

497,059 high-quality genotyped SNPs

Whole-genome
genotype imputation

8,530,563 genetic variants for GWAS

Significant loci
($P < 5.0 \times 10^{-8}$)

Functional characterization by
in silico and *in vitro* analyses

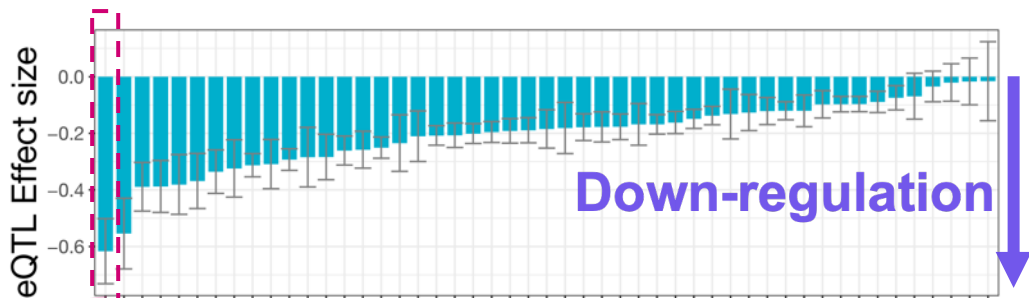
Replication Analysis

 99 cases  1,026 controls

Replication analysis with independent data

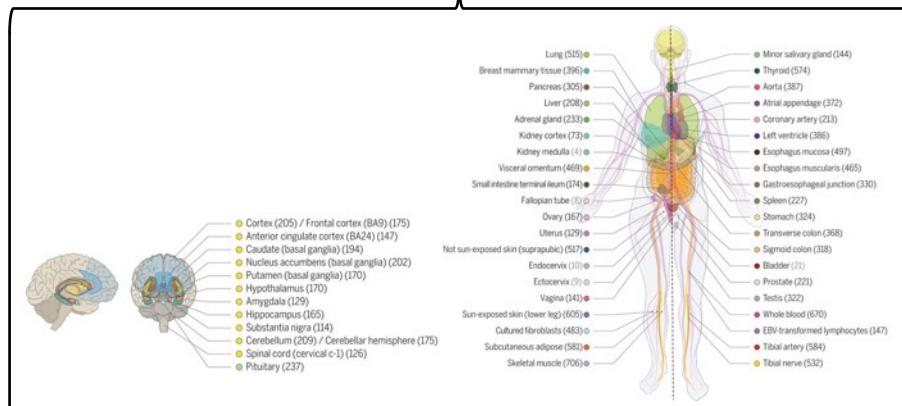
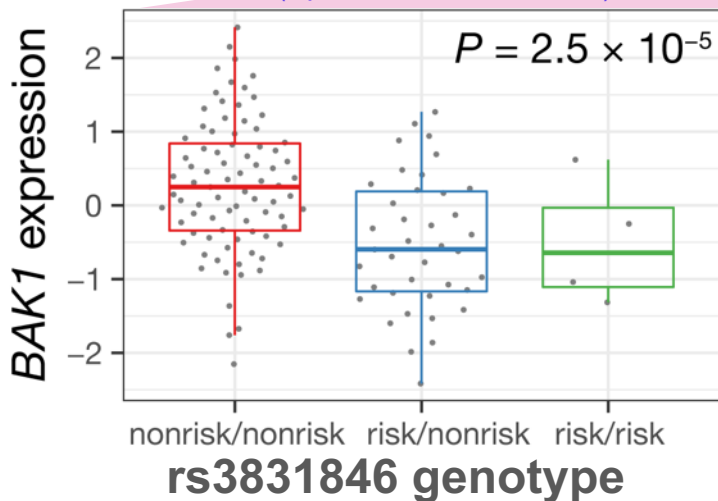
rs3831846 has eQTL effect on *BAK1*

Expression quantitative trait locus (eQTL) analysis revealed **the risk allele down-regulating *BAK1*** in a broad range of tissues.



49 tissues
in the GTEx project

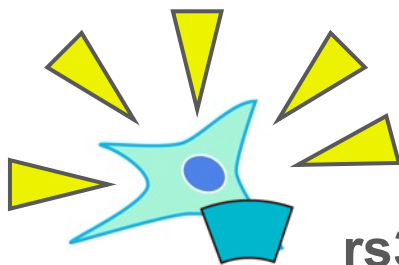
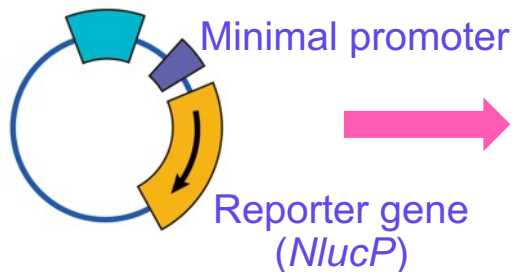
Brain (Spinal cord cervical c-1)



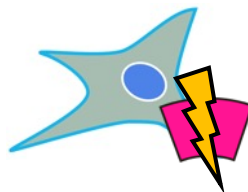
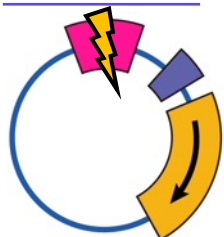
rs3831846 reduces *BAK1* enhancer activity

In vitro reporter assay experimentally validated
the risk allele of rs3831846 attenuating the *BAK1* enhancer activity.

Enhancer sequence
with **Non-risk allele**

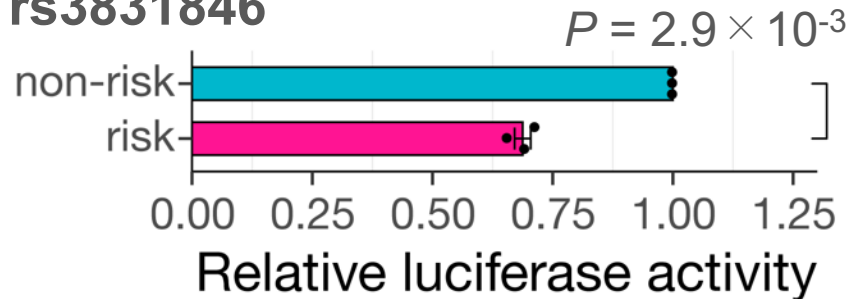


Enhancer sequence
with **Risk allele**



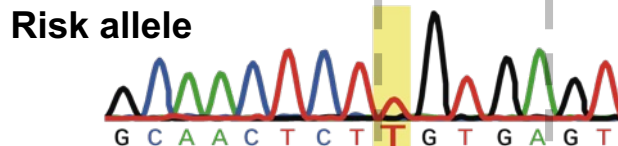
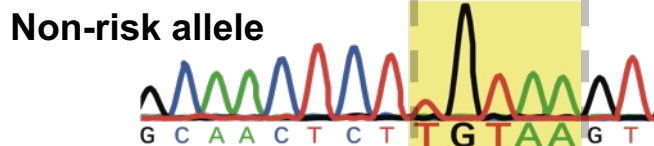
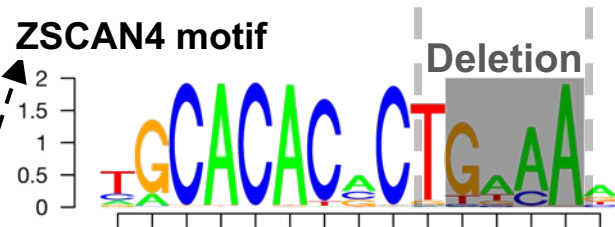
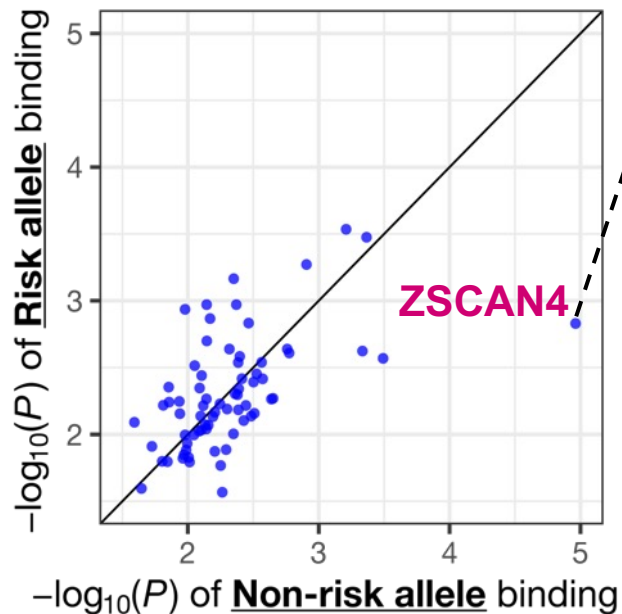
transfected 293T cells

rs3831846

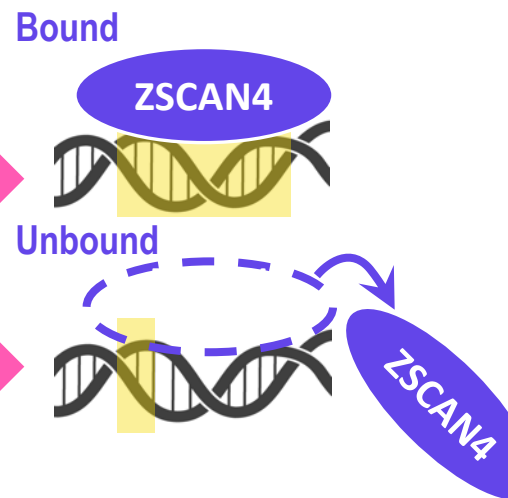


Disrupting Impact on TF Binding Motif

We evaluated transcription factor binding score using the JASPAR database. The deletion has **disrupting impact on the binding motif in the enhancer.**



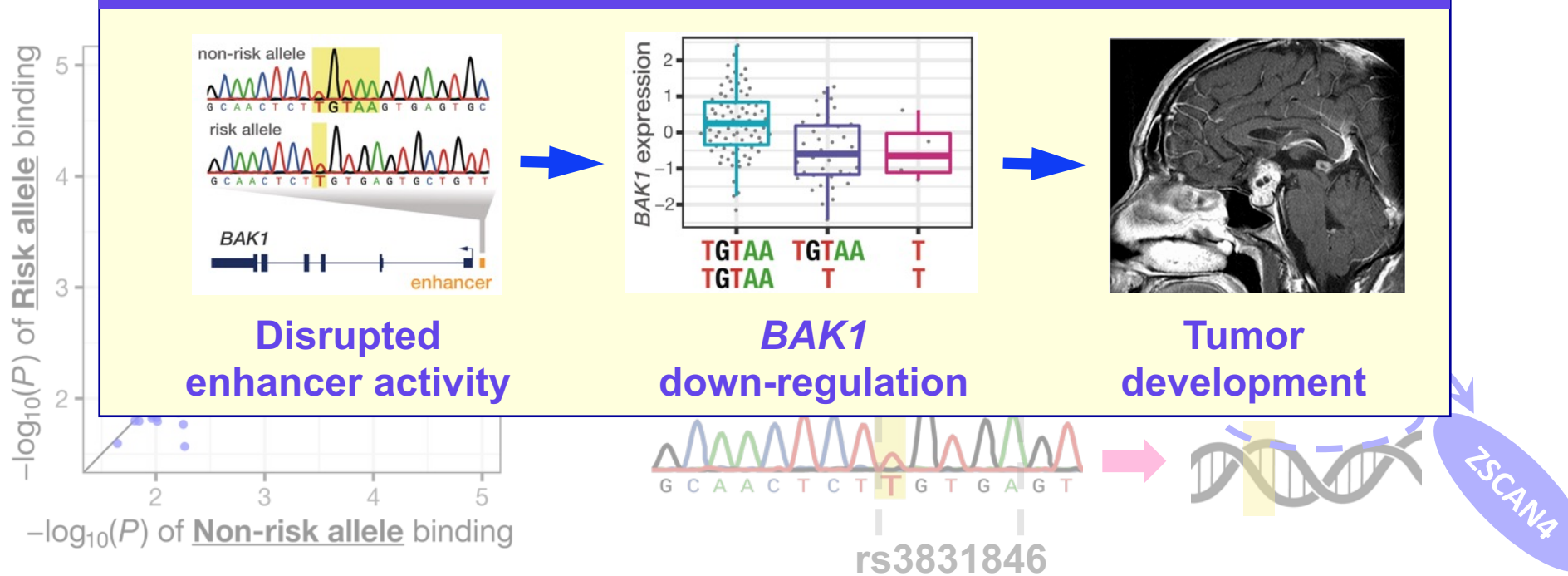
rs3831846



Disrupting Impact on TF Binding Motif

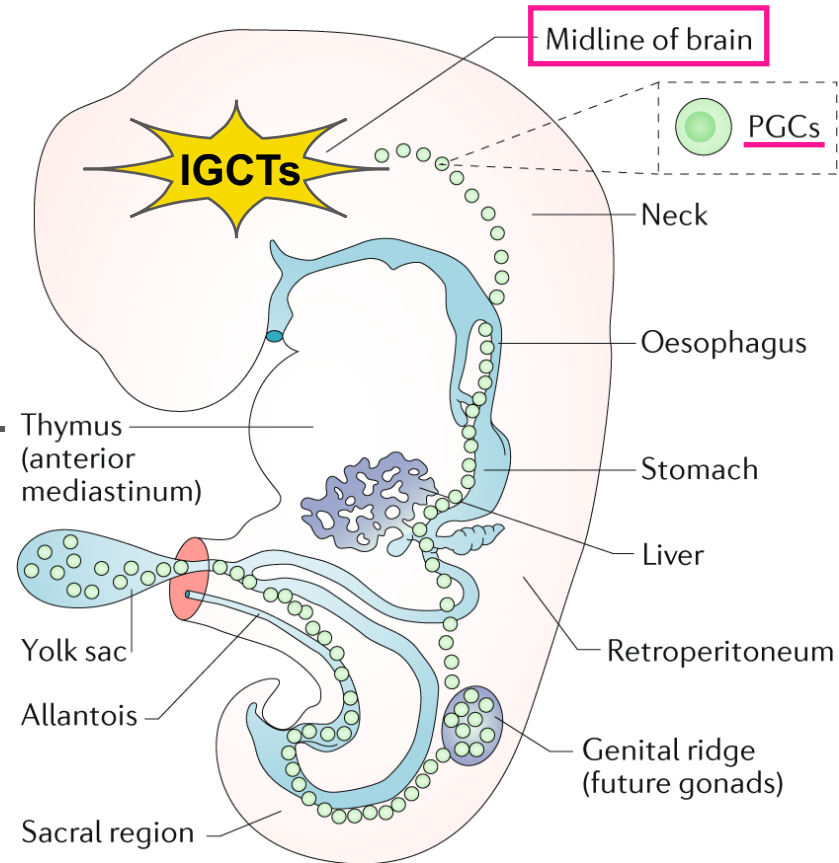
We evaluated transcription factor binding score using the JASPAR database.

The **Molecular mechanism behind the risk association** cer.



BAK1 in IGCTs Etiology (from embryology)

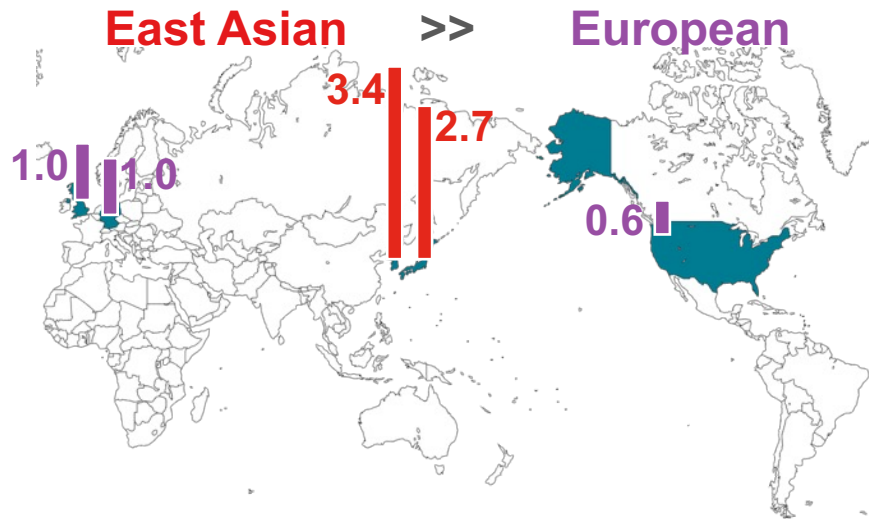
- IGCTs are considered originated from the embryonic precursor of sperm and egg (**primordial germ cells: PGCs**).
- During embryogenesis, some PGCs can mis-migrate into the brain, **physiologically removed by apoptosis**.
- The down-regulation of *BAK1* may allow **those mis-migrated PGCs to escape the removal and form IGCTs**.



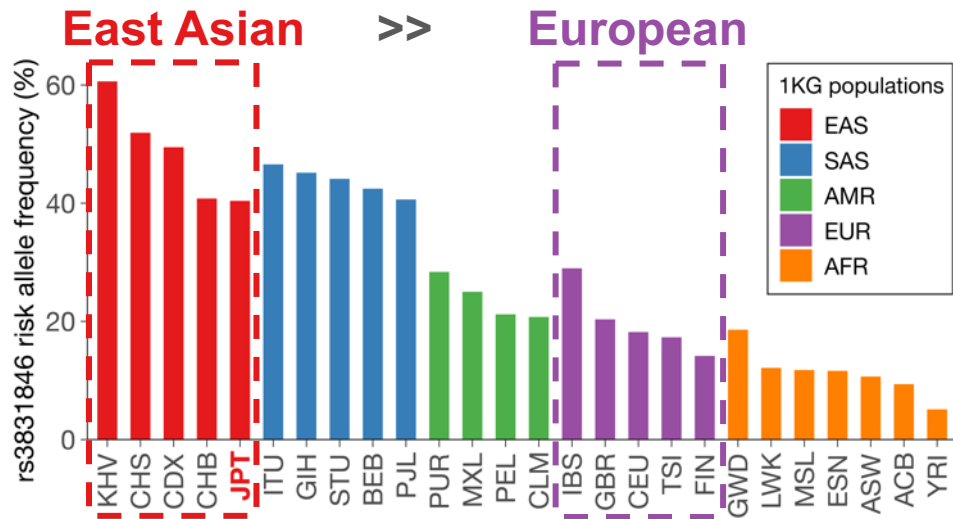
BAK1 rs3831846 in IGCs Epidemiology

- The risk allele frequency of rs3831846 is **48.5% in East Asian** and **19.9% in European**.
- The proportion of the risk allele carriers may contribute to the geographical difference in incidence.

Incidence (/million/year)

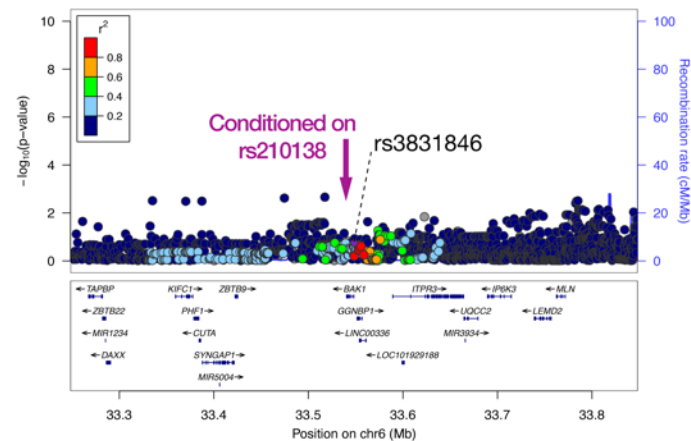
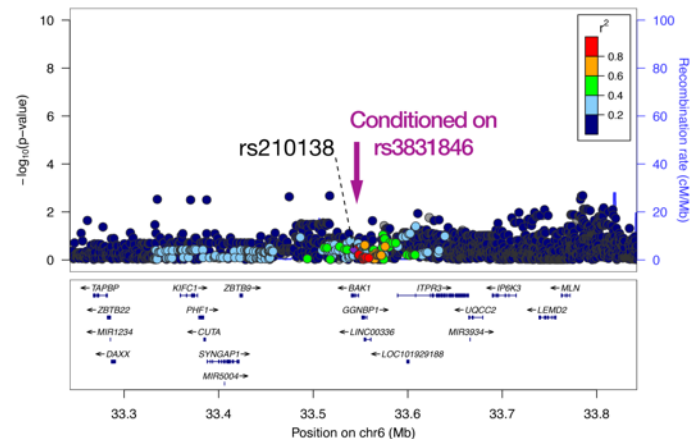
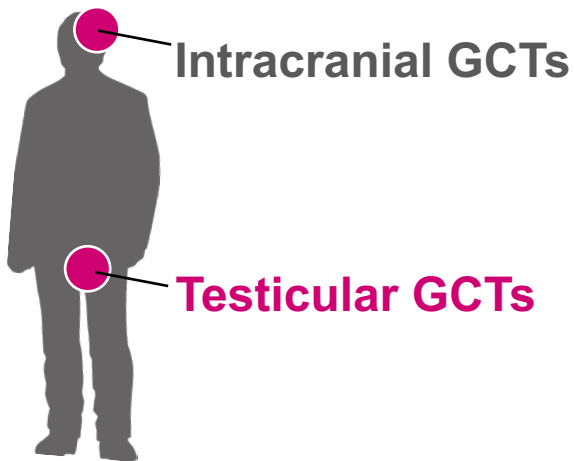


Risk allele frequency



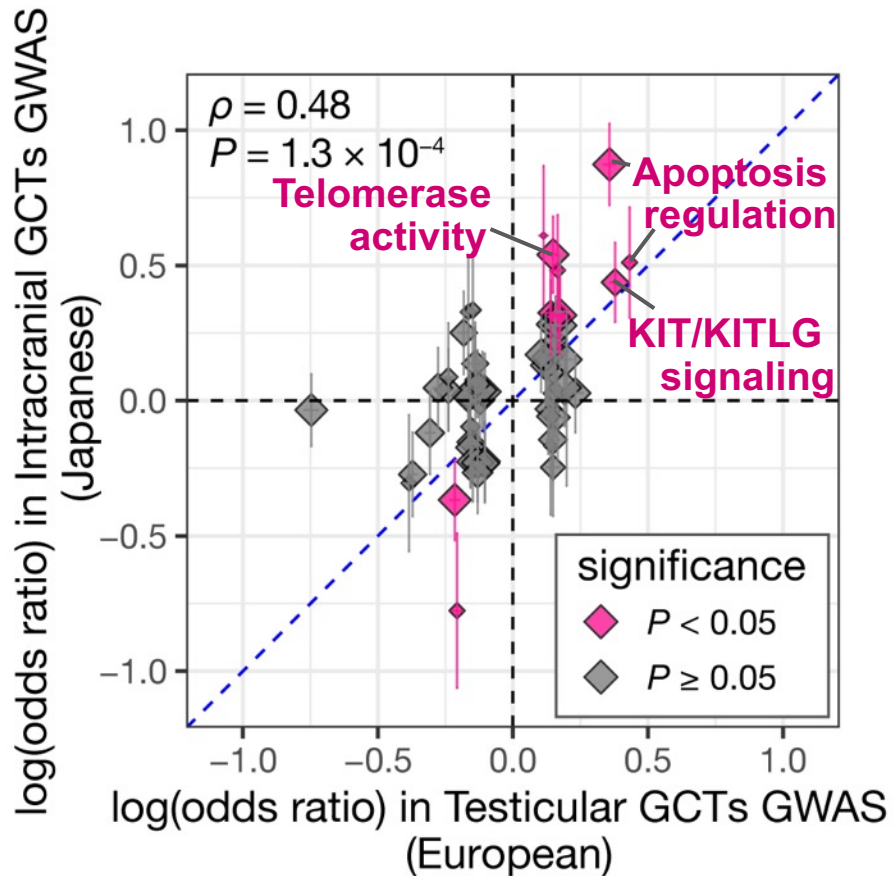
Shared Genetic Basis with Testicular GCTs

- **Testicular GCTs (TGCTs)** are major testicular cancer histologically similar to IGCTs.
- Rs3831846 is in strong linkage disequilibrium with TGCTs risk SNP rs210138 ($r^2 = 0.98$ in both European and East Asian).



Shared Genetic Basis with Testicular GCTs

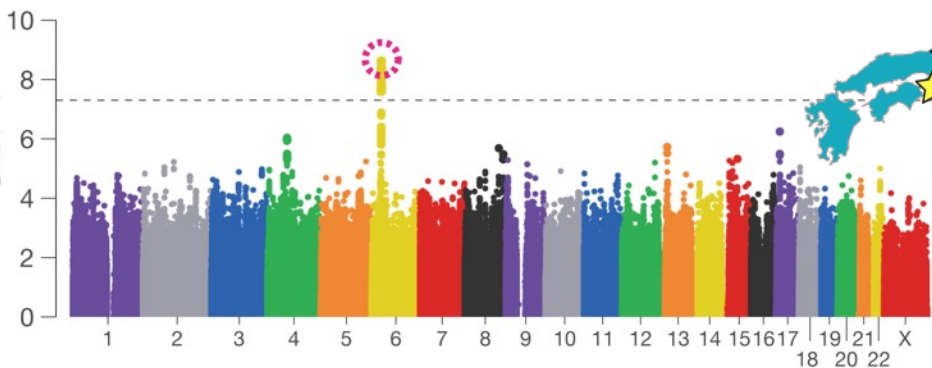
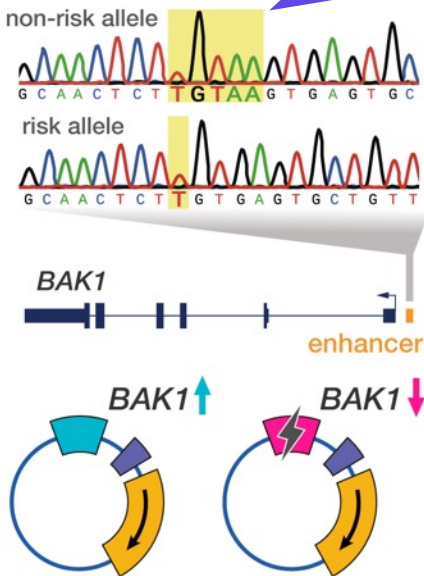
- Among 57 TGCTs-risk loci available in our GWAS, **11 loci showed $P < 0.05$ for IGCTs risk**, with concordant risk alleles.
- These significant loci were implicated in multiple biological pathways, such as:
 - *BAK1* and *SPRY4* from KIT/KITLG signaling
 - *CLPTM1L* from apoptosis regulation
 - *PITX1* from telomerase activity



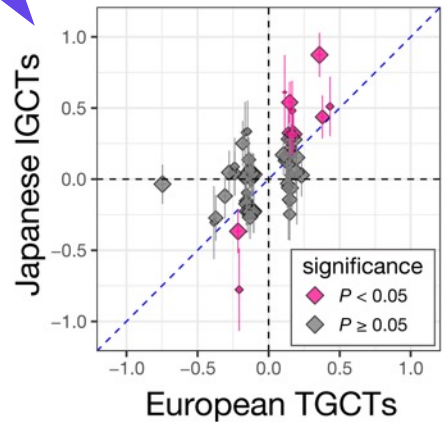
Summary

The first IGCTs GWAS

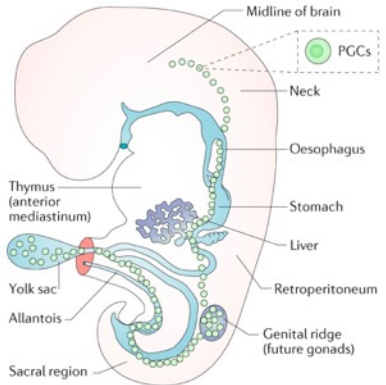
Molecular mechanism



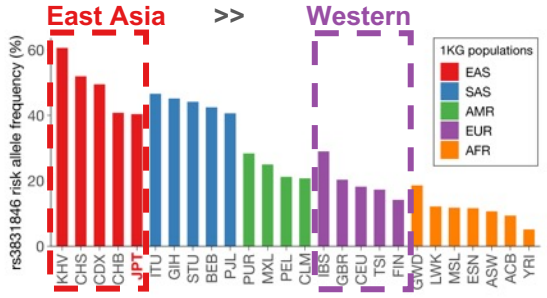
Shared genetic basis



Etiology



Epidemiology



Acknowledgements

We sincerely thank all the participants involved in this study.

Graduate School of Medicine, the University of Tokyo

Osaka University Graduate School of Medicine

RIKEN Center for Integrative Medical Sciences

Yukinori Okada, Qingbo Wang, Ken Suzuki, Saori Sakaue, Toshihiro Kishikawa, Kenichi Yamamoto, Tatsuhiko Naito, Tsuyoshi Deguchi, Tomoko Nakanishi, Yuya Shirai, Kaori Odomari, Shinichi Namba, Yoshihiko Tomofuji, Noa Sasa, Ryuya Edahiro, Tomohiro Yata, Goh Sato, Takafumi Ojima, Yuji Yamamoto, Hiroyuki Ueda, Yuwa Oka, Yosuke Ogawa, Yohei Tsukazaki Chifune Kai, Genta Nagao

National Center for Child Health and Development

Keita Terashima, Nobutaka Kiyokawa, Hideki Ogiwara

National Cancer Center Research Institute

Koichi Ichimura, Tatsuya Ozawa, Yoshiko Nakano, Yui Kimura

Kyoto University Graduate School of Medicine

Fumihiko Matsuda, Meiko Takahashi

Graduate School of Medicine, Juntendo University

Arata Tomiyama, Yuko Matsushita, Takashi Fujii

Saitama Medical University International Medical Center

Ryo Nishikawa, Tomonari Suzuki

Faculty of Medicine, Hokkaido University

Shigeru Yamaguchi

Keio University School of Medicine

Hikaru Sasaki, Tomoru Miwa

Shizuoka Children's Hospital

Masashi Kitagawa

Graduate School of Medicine, Gifu University

Naoyuki Ohe

Wakayama Medical University School of Medicine

Junya Fukai

Hyogo Prefectural Kobe Children's Hospital

Atsufumi Kawamura

Faculty of Medicine, the University of Tokyo

Satoru Miyawaki