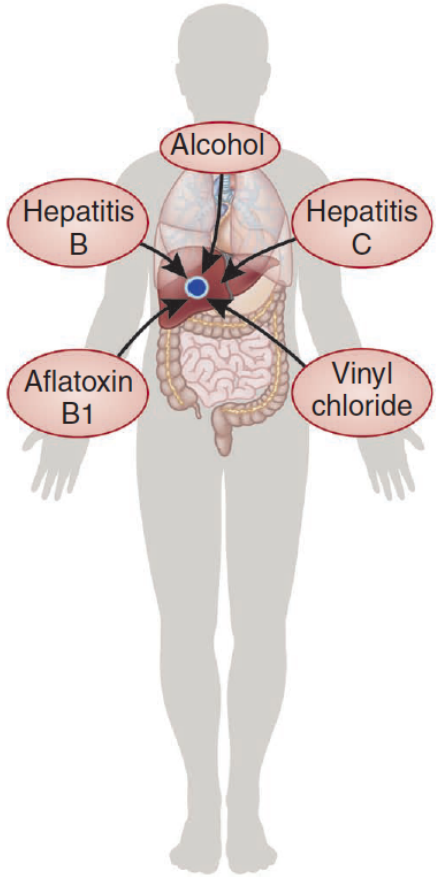
# Whole-genome mutational landscape and characterization of noncoding and structural mutations in liver cancer

Akihiro Fujimoto[1], Mayuko Furuta[1], Yasushi Totoki[2], Tatsuhiko Tsunoda[1], Mamoru Kato[2], Atsushi Ono[4], Hiroki Yamaue[3], Kazuaki Chayama[1,4], Satoru Miyano[5], Hiroyuki Aburatani[6], Tatsuhiro Shibata[2], Hidewaki Nakagawa[1]

(1; IMS, RIKEN, 2; Natl. Cancer Ctr. Res. Inst., 3; Wakayama Med. Univ., 4; Hiroshima Univ. Sch. of Med., 5; Inst. of Med. Sci., The Univ. of Tokyo, 6; RCAST, The Univ. of Tokyo)

# Liver cancer

✓**Liver cancer is the third leading cancer-related cause of death and the seventh most common form of cancer in the world.**

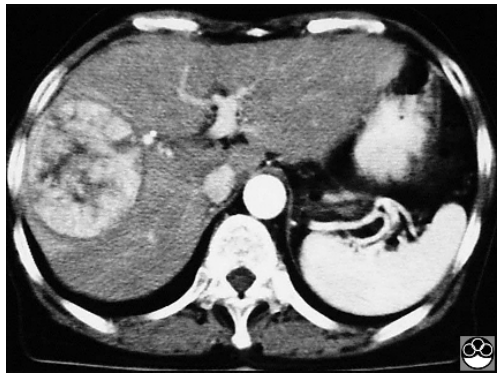



✓**The mortality rate of liver cancer is high in East Asia and Africa, and is also increasing in western countries.**

✓**One of the major risk factors is virus infection;**

**Hepatitis B virus (HBV): Asia, Africa**

**Hepatitis C virus (HCV): Japan (>70% of HCC)**



✓ **Large-scale whole genome sequencing study has not yet been performed for liver cancer.**

Figure from Zhang *Nat. Genet.* **44**, 1075-7 (2012)

# Topics

- Whole genome sequencing of 300 liver cancer genomes

  Fujimoto *et al*. **Nature Genetics** (2016)

- Circulating tumor DNA analysis for liver cancers

  Ono *et al*. **Cellular and Molecular Gastroenterology and Hepatology** (2015)

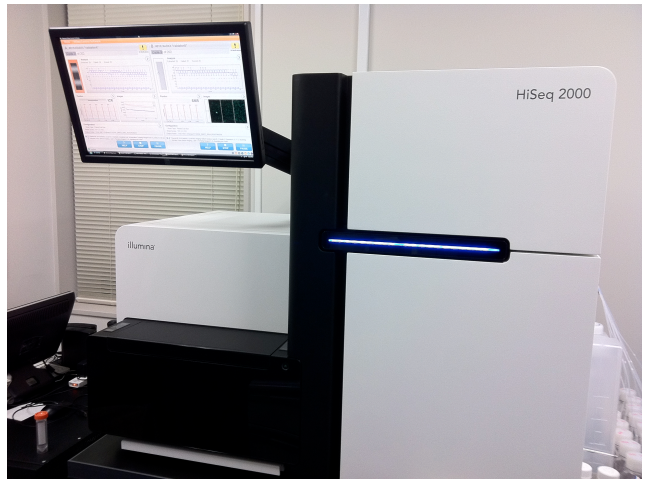
# Whole genome sequencing of 300 liver cancer genomes

Fujimoto *et al*. Nature Genetics (2016)

# Sample and sequencing

- **Sample**
  - 300 liver cancers
  - 268 hepatocellular carcinomas (HCCs), 24 intrahepatic cholangiocarcinomas (ICCs), 8 cHCC/ICCs (combined type)
  - 159 HCV, 82 HBV, 55 NBNC and 4 HBV/HCV
  - Whole genome sequencing (WGS) and RNA-sequencing

- **Sequencing**
  - HiSeq2000
  - Paired-end method
  - Library size; 500bp
  - Read length; 100bp

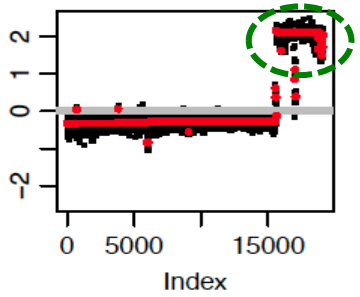


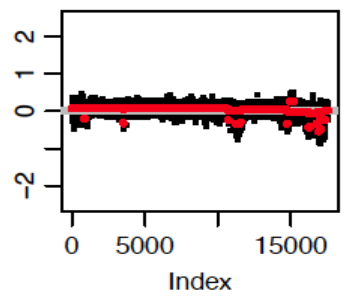
# Somatic mutation calling

**Point mutation and short indels**

- Based on number and frequency of mutant alleles
- False positive rate of point mutation < 3 % in CDS region
- False positive rate of indel < 10 %

Amplification



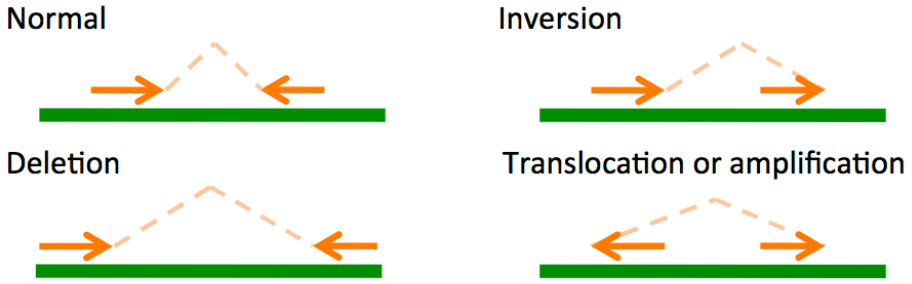



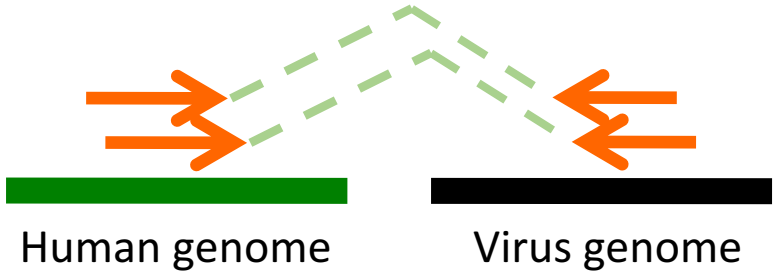

**Copy number alternation**

- Ratio of depth of coverage was analyzed by DNAcopy[1]

**Structural variations (STV) and virus integration**

- Based on orientation and distance between read-pairs
- Realignment and comparison with other normal samples were performed to exclude false positives




1; Andersson *et al*. Bioinformatics (2008)

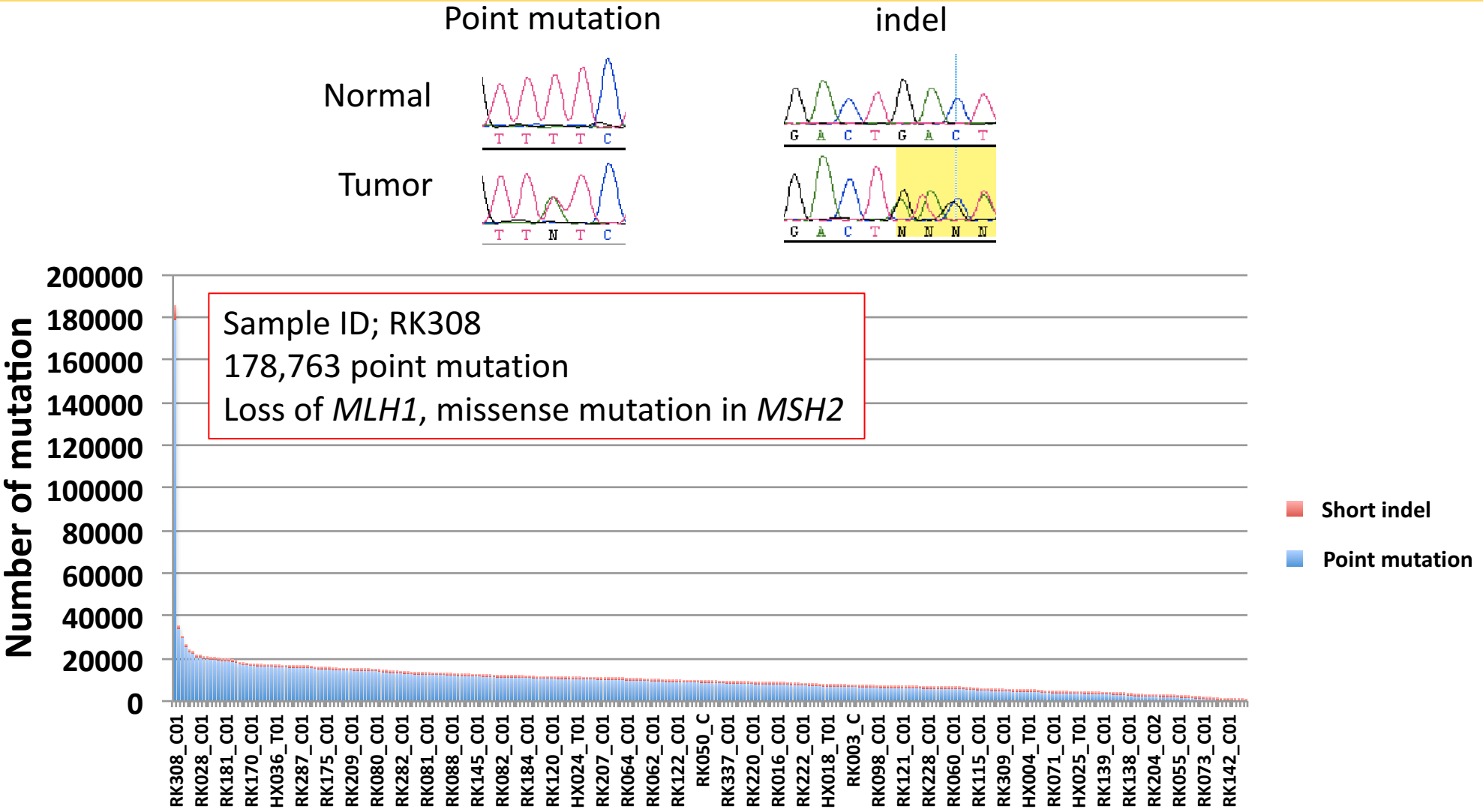
# Contents

- Pattern of mutation
  - Point mutation and short indel
  - Structural variation
  - Integration of virus into human genome
- Driver genes and noncoding driver regions
  - Significantly mutated genes
  - Significantly mutated noncoding regions
- Structural variations
  - Structural variations
  - Structural variations and gene expression
  - Mutation and gene expression in *TERT*

# Contents

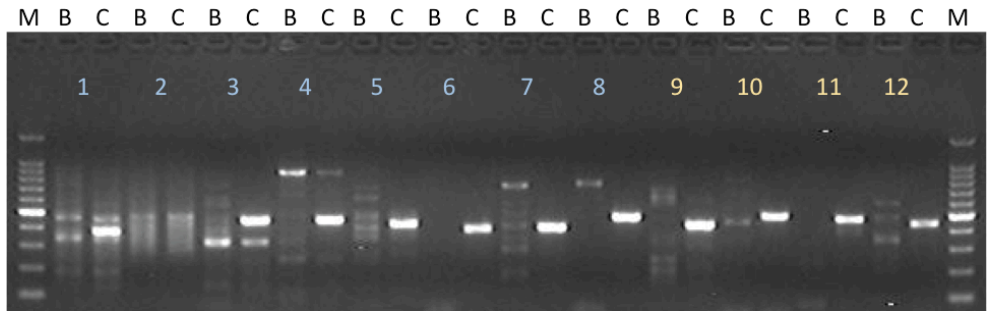
- Pattern of mutation
  - Point mutation and short indel
  - Structural variation
  - Integration of virus into human genome
- Driver genes and noncoding driver regions
  - Significantly mutated genes
  - Significantly mutated noncoding regions
- Structural variations
  - Structural variations
  - Structural variations and gene expression
  - Mutation and gene expression in *TERT*

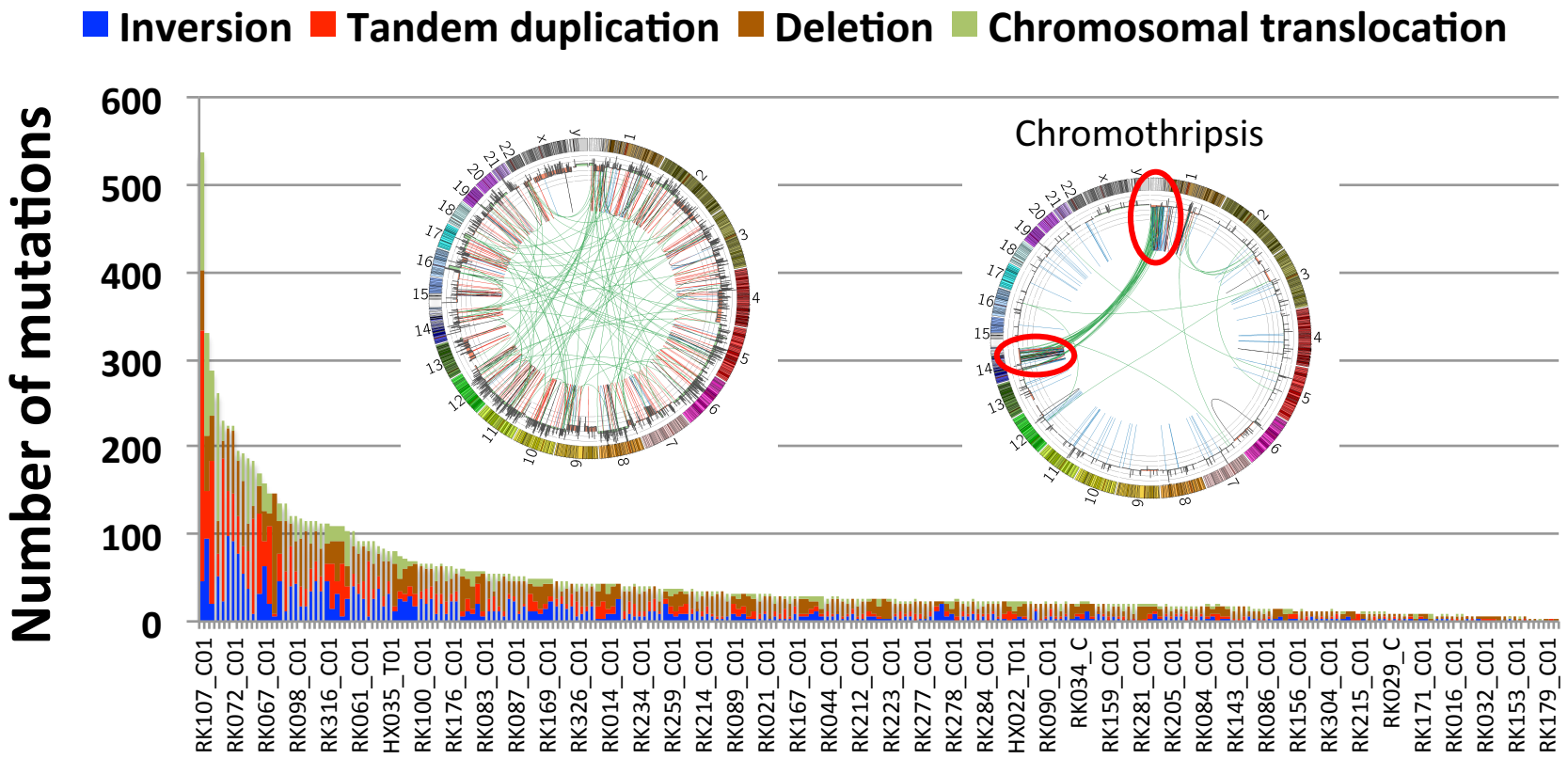
# Point mutation and short indel




Number of point mutation; 8,777, short indel; 239 (median)

# Structural variation (STV)



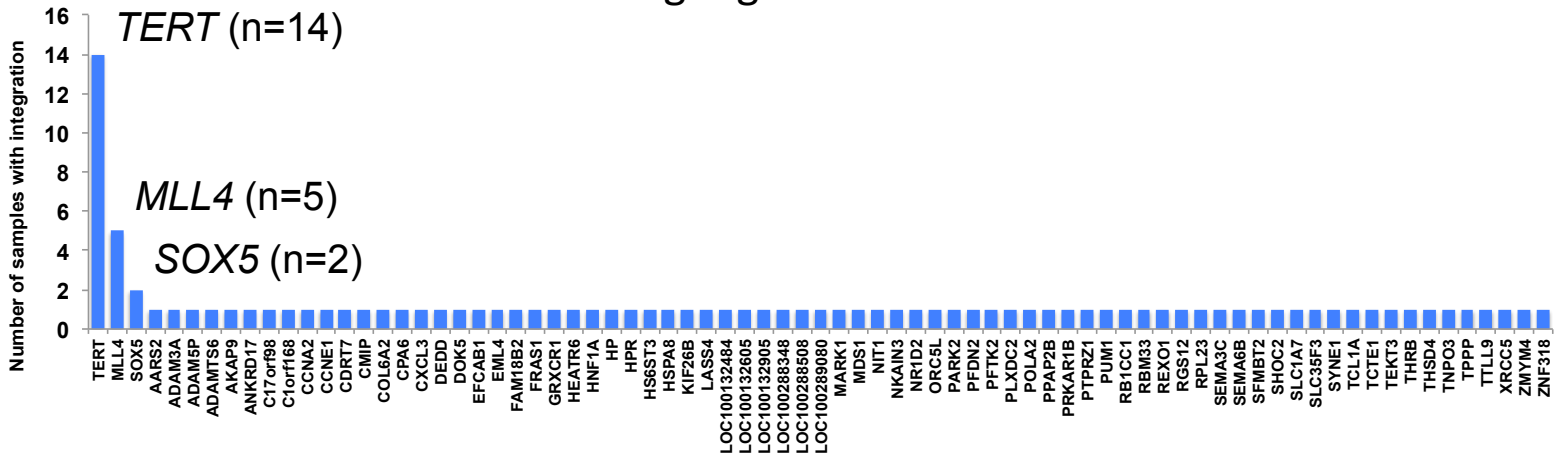




Number of structural variation; 40.6

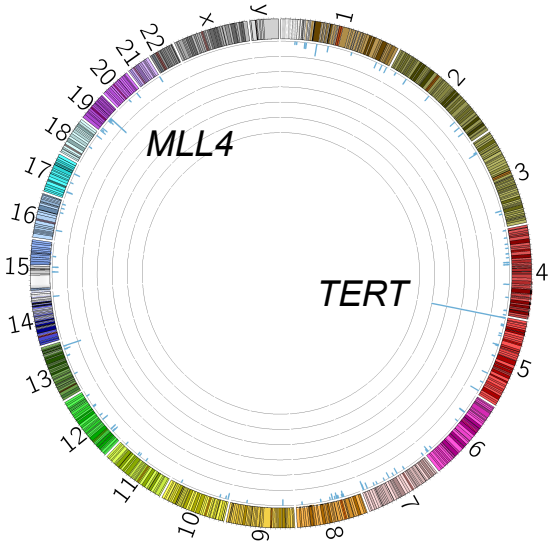
# Target gene of HBV integration

223 HBV integration events in 64 samples were identified.

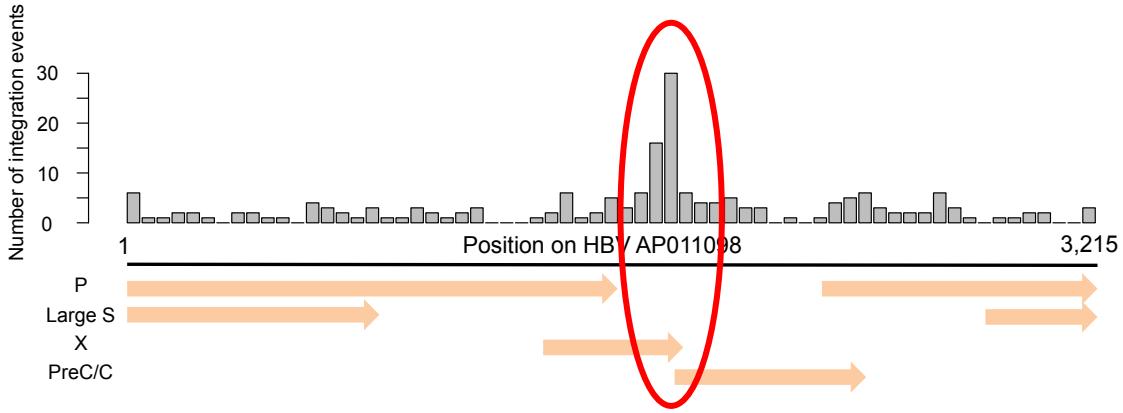
Target gene




HBV integration sites in human genome




HBV integration sites in HBV genome




Downstream of X gene

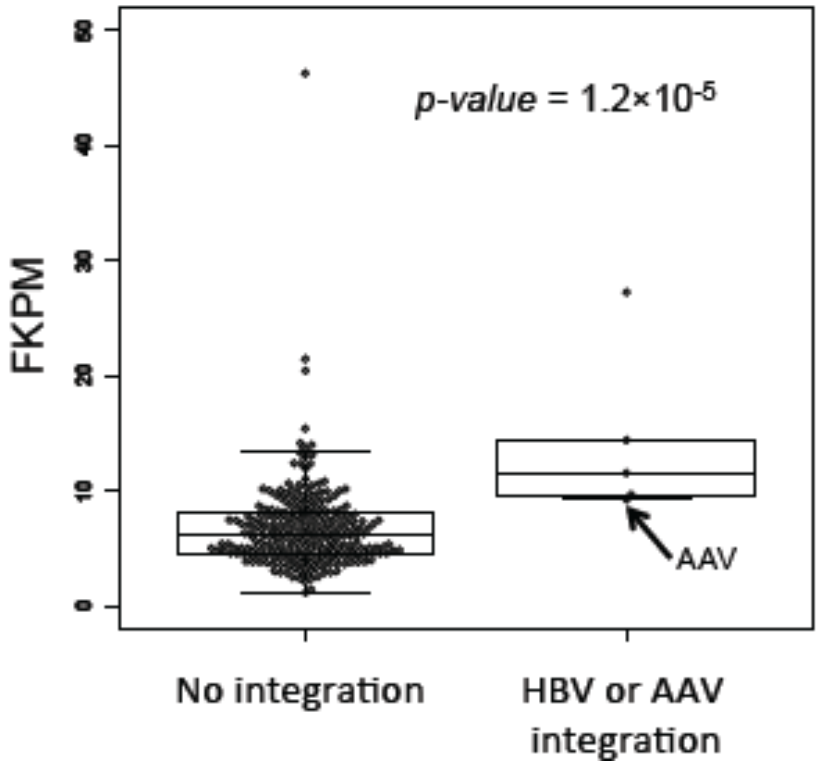
# Identification of adeno-associated virus (AAV)

- *De novo* assembly of the unmapped reads from RNA-seq generated some long contigs that aligned to adeno-associated virus (AAV) in two liver cancer samples.

Integrated genes

| | AAV genome | | Human genome | | | |
|---|---|---|---|---|---|---|
| Sample | AAV sequence | Position | Chr | Position | Number of support read pair | Gene |
| HX032_T01 | NC_001401 | 4386 | 19 | 30304517 | 32 | *CCNE1* intron |
| HX032_T01 | NC_001401 | 4600 | 19 | 30304511 | 17 | *CCNE1* intron |
| RK112_C01 | NC_001401 | 3320 | 19 | 36213603 | 1 | *MLL4* intron |
| RK112_C01 | NC_001401 | 4438 | 19 | 36213955 | 1 | *MLL4* intron |
| RK236_C01 | NC_001401 | 2982 | 5 | 18813398 | 3 | |
| RK236_C01 | NC_001401 | 82 | 5 | 18813512 | 13 | |

Expression level of *MLL4* gene

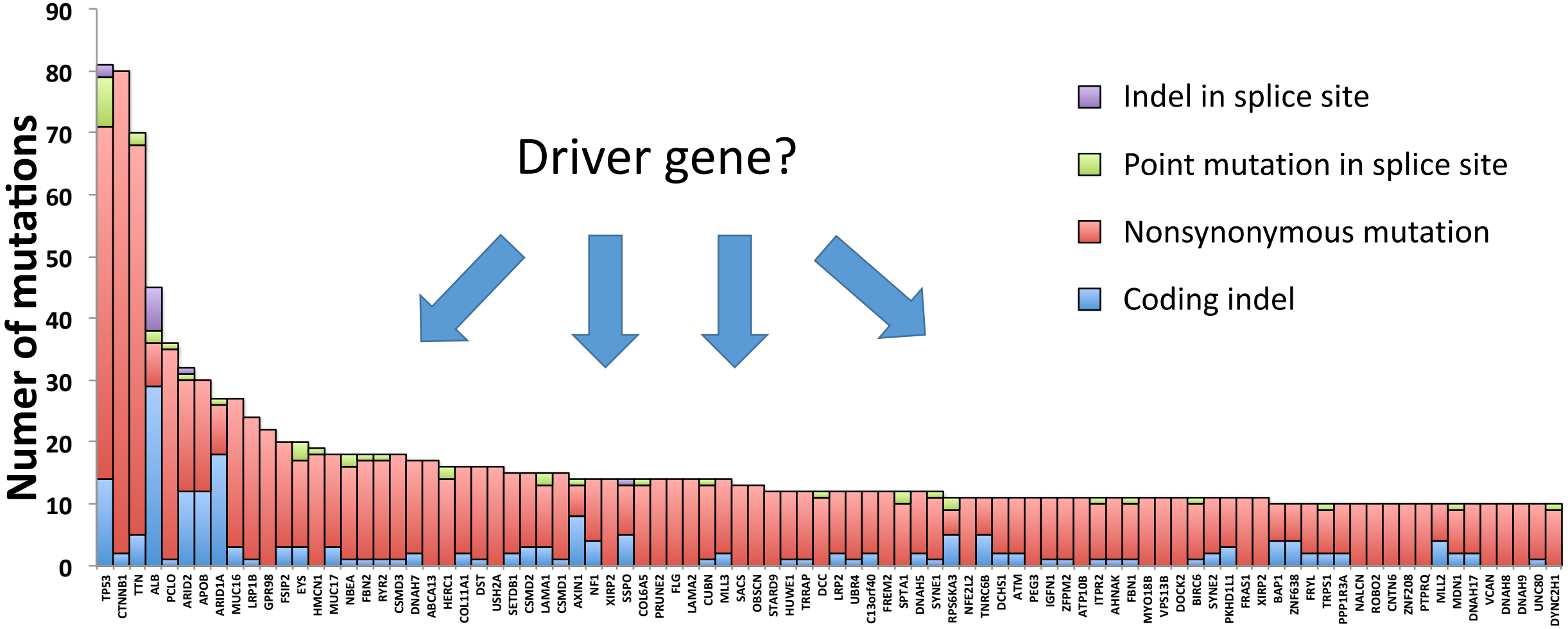



- Integration events of AAV into *MLL4* and *CCNE1* gene were identified. Expression level of *MLL4* gene in the integrated samples were higher than that of others. There results are consistent with a recent study (Nault *et al.* Nat Genet (2015))。

# Contents

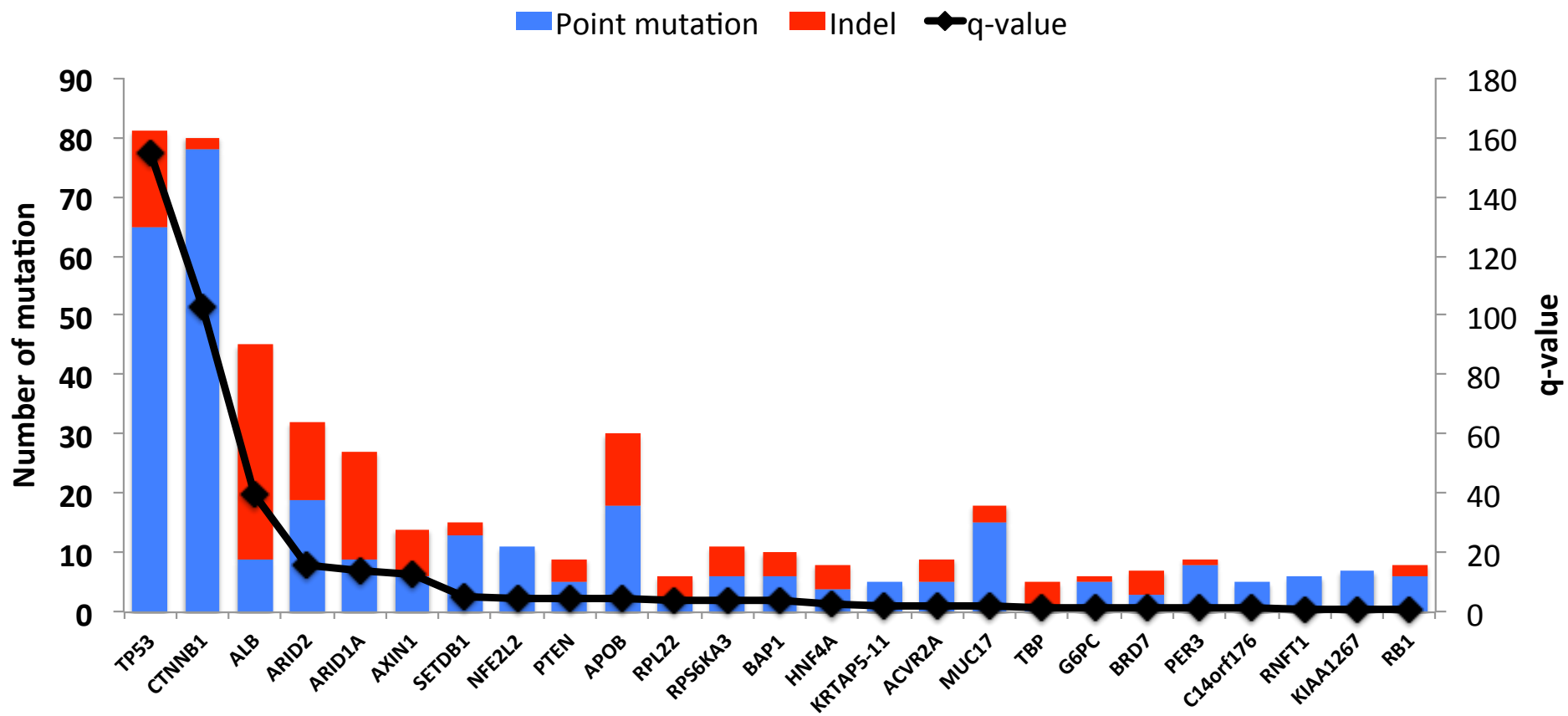
- Pattern of mutation
  - Point mutation and short indel
  - Structural variation
  - Integration of virus into human genome
- Driver genes and noncoding driver regions
  - Significantly mutated genes
  - Significantly mutated noncoding regions
- Structural variations
  - Structural variations
  - Structural variations and gene expression
  - Mutation and gene expression in *TERT*

# Protein alternating mutations（300 liver cancers）



# Identification of significantly mutated genes (1)

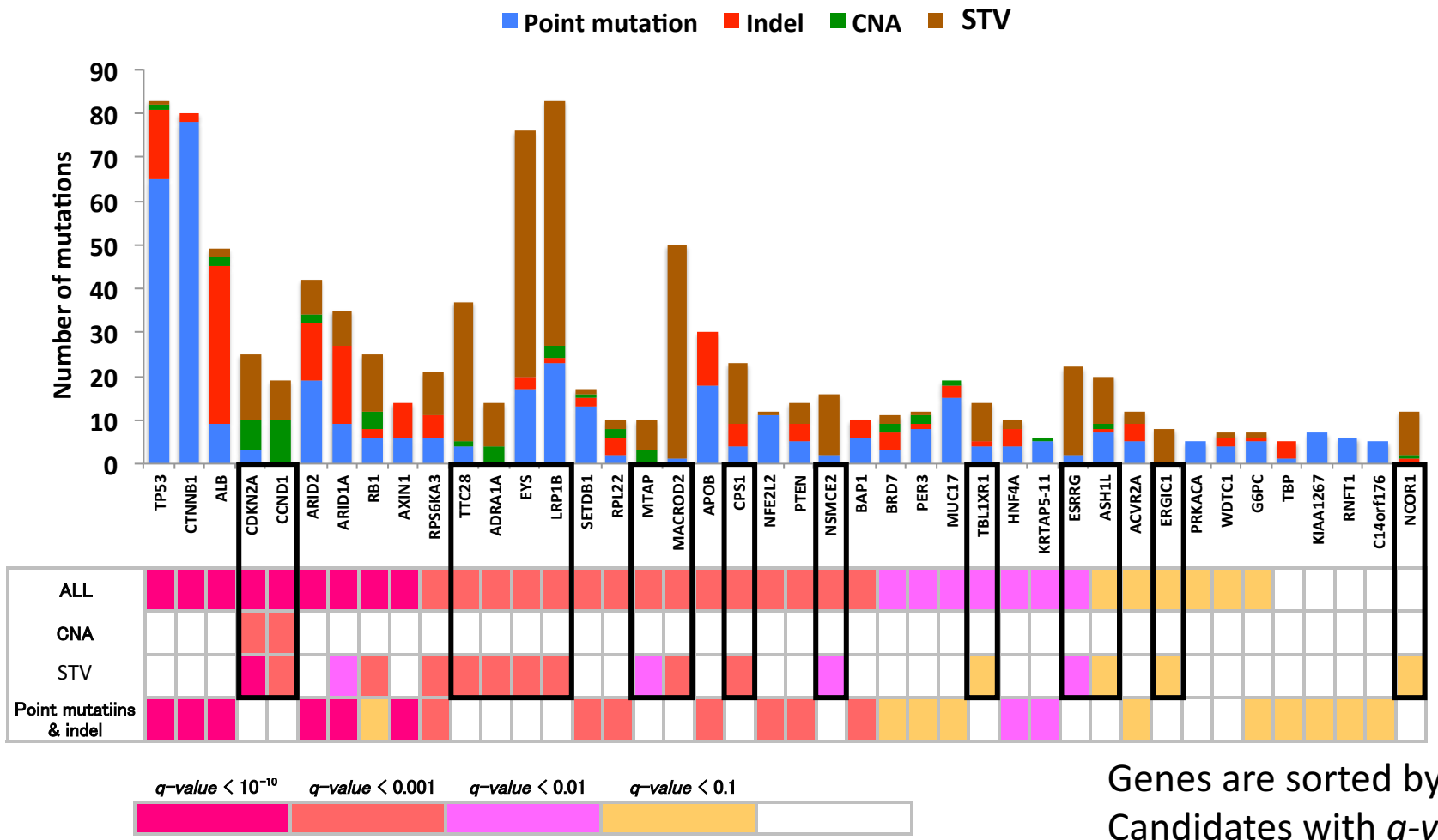
(1) Estimate noncoding mutation rate within 1Mbp region
(2) Calculate the expected number of mutations for each gene ((length) × (mutation rate))
(3) Calculate probability that the observed number of mutation ≥ expected number of mutations under the Poisson distribution




- Twenty-five genes, including *TP53*, *CTNNB1*, *ARID2*, *ARID1A*, *RB1*, *AXIN1*, *RPS6KA3*, *SETDB1*, *NFE2L2*, *BAP1*, and *HNF4A*, had a significantly large number of protein-altering mutations.

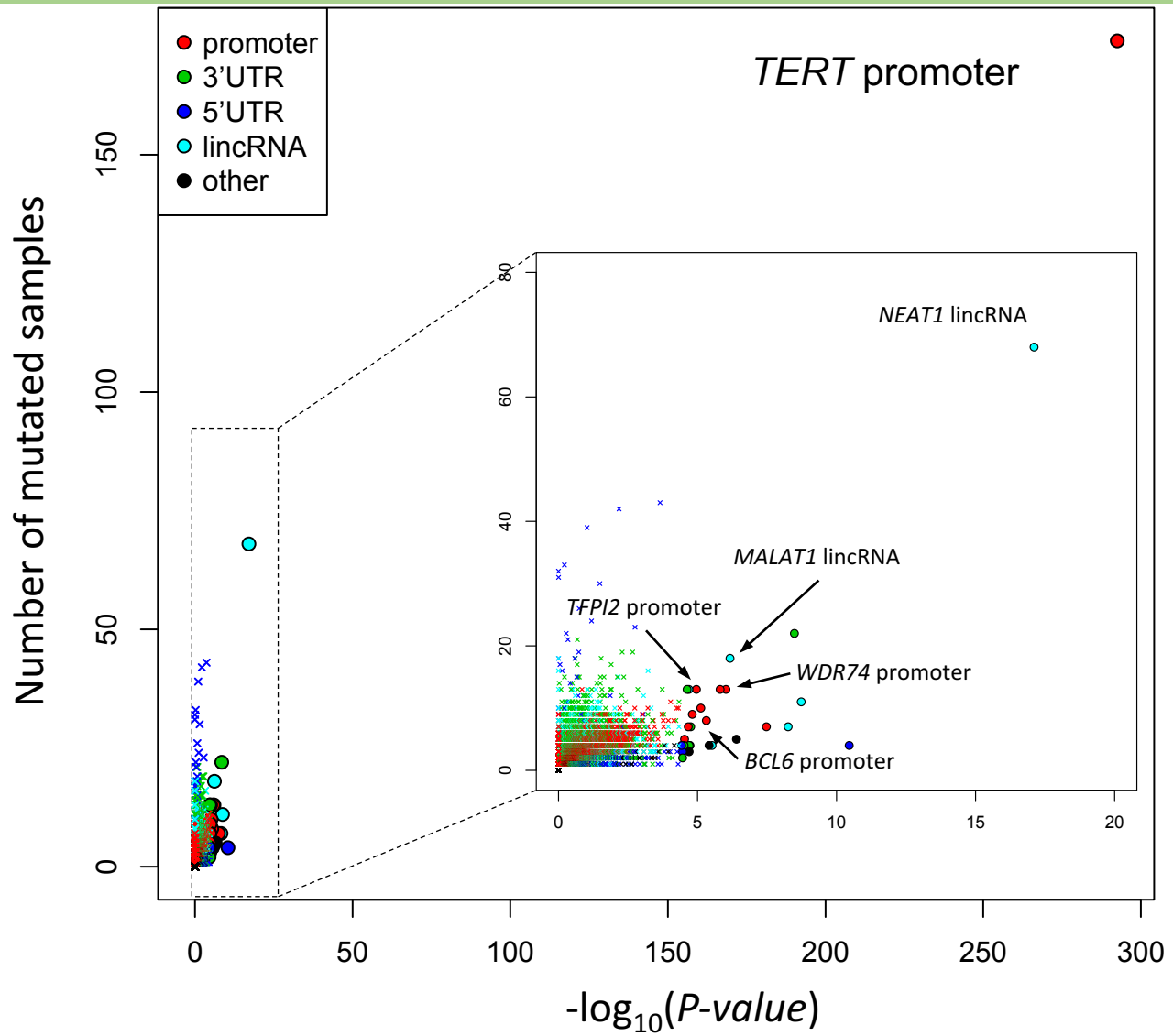
# Identification of significantly mutated genes (2)

- We then combined the information on point mutations, indels, STVs, and CNAs and tested the significance of the number of mutations for each gene.
- Thirty-eight genes had a significantly large number of mutations.
- Genes with the terms ‘chromatin regulator’ and ‘regulation of cell cycle’ were significantly enriched.



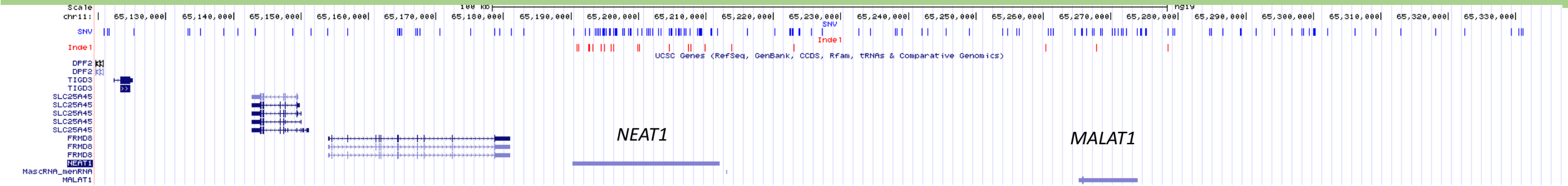

Genes are sorted by *q-value*.
Candidates with *q-value* < 0.1 are shown.

# Identification of significantly mutated noncoding elements (GENCODE annotation)




- In addition to the known *TERT* promoter, six long intergenic noncoding RNA (lincRNA) genes (*NEAT1* and *MALAT1*), ten promoter (*TFPI2*, *MED16*, and *WDR74*), and nine UTRs (*BCL6* and *AFF4*) were identified as regions with a significantly large number of mutations.

# lincRNA (*NEAT1* and *MALAT1*)




- *NEAT1* (Nuclear Enriched Abundant Transcript 1)

  *NEAT1* has an essential role in constructing a subnuclear structure, paraspeckle[1].

  is differentially expressed in several cancers[2,3,4,5]

  is induced by hypoxia and promote cell proliferation and invasion.

  promotes cell proliferation and invasion through regulating miRNAs[6,7,8,9,10].

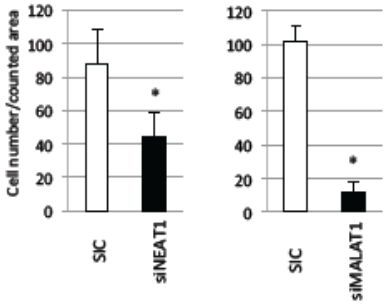
  is induced by TP53 and modulate replication stress response and chemosensitivity[11].

- *MALAT1* (Metastasis-associated lung adenocarcinoma transcript 1)

  Expression level of *MALAT1* was associated with prognosis in liver cancer[8].

- Knockdown of *NEAT1* or *MALAT1* in HepG2 cells decreased cell invasion

Invasion assay



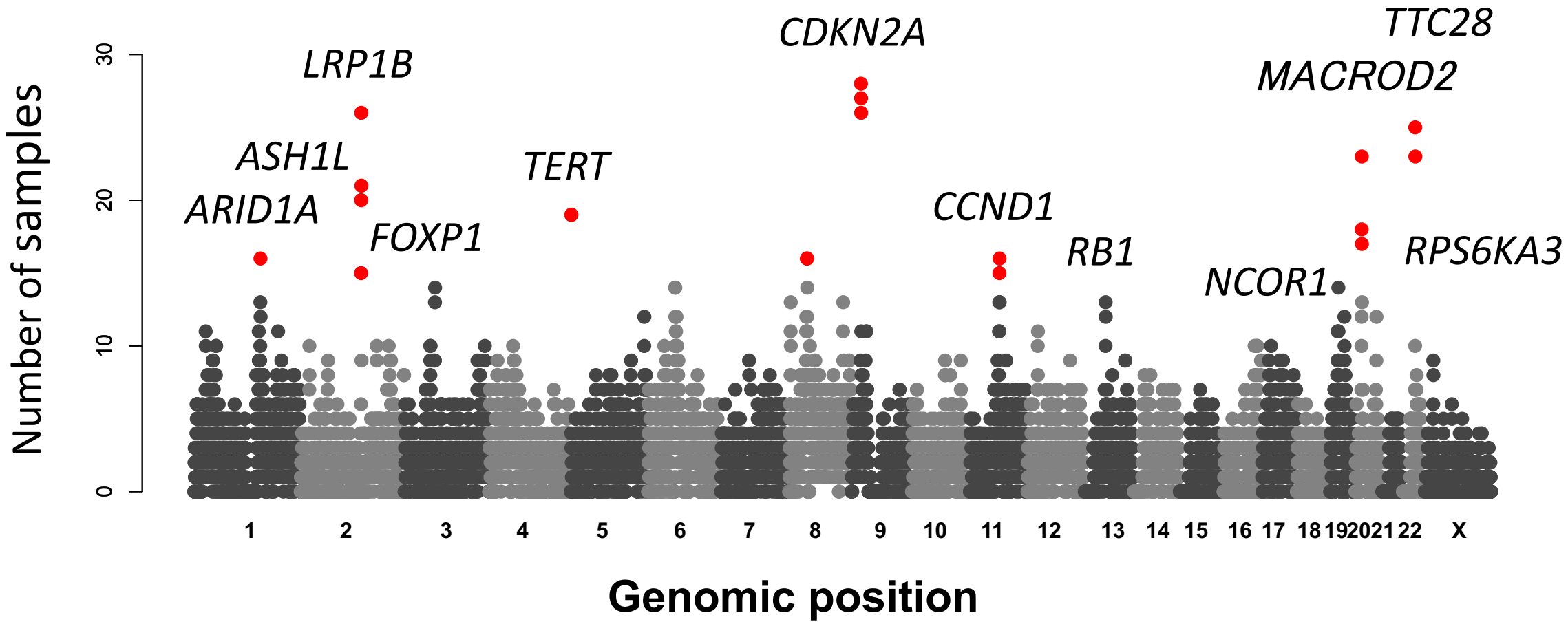

1; Hirose *et al.* Mol. Biol. Cell (2014), 2; Li *et al.* Oncotarget (2015), 3; Wu et al. Mol. Cancer (2016), 4; Chakravarty et al. Nat Comms (2014), 5 Choudhry et al Oncogene (2015). 6; Wang et al. JECCR (2016), 7; Cao et al. Am J Cancer Res (2016), 8; Huang et al. BBRC (2016), 9; Ke et al. GRSB (2016), 10; Lai *et al*. Med Oncol (2012), 11 Adriaens et al. Net Med (2016)

# Contents

- Pattern of mutation
  - Point mutation and short indel
  - Structural variation
  - Integration of virus into human genome
- Driver genes and noncoding driver regions
  - Significantly mutated genes
  - Significantly mutated noncoding regions
- Structural variations
  - Structural variations
  - Structural variations and gene expression
  - Mutation and gene expression in *TERT*

# Distribution of structural variation (STV)

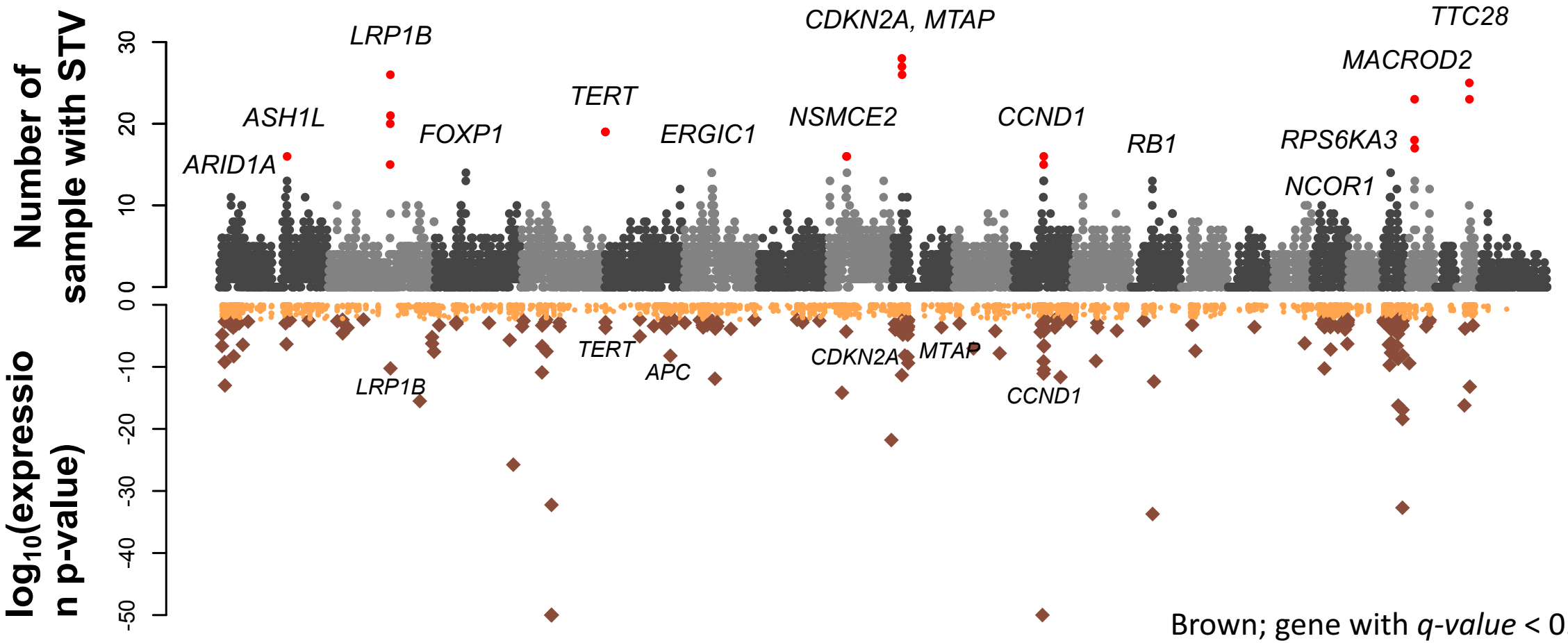
Number of samples with an STV breakpoint within 500-kb bins.




Red; Bins containing STV breakpoints in ≥15 samples (5%)

# STVs and gene expression

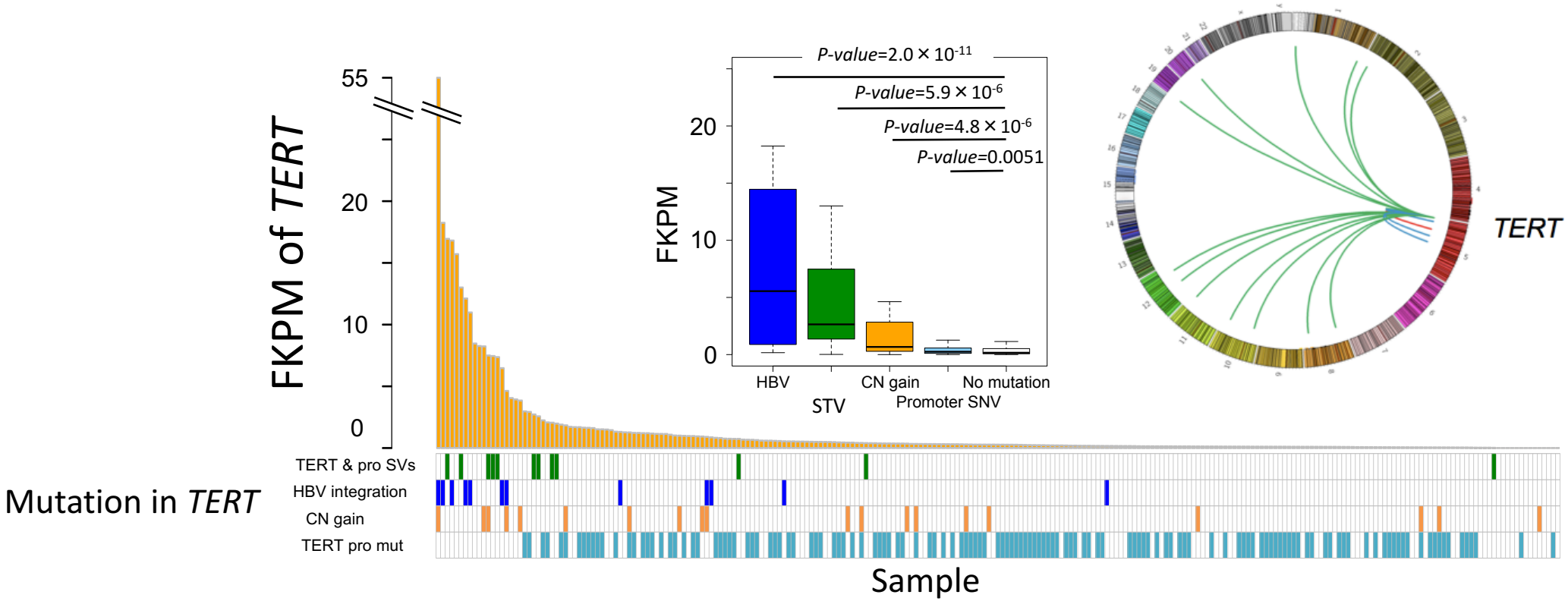
- We selected 4,940 genes that had STVs within 500 kb of their transcription start or end sites in ≥6 samples (2%) and compared the number of RNA-seq reads mapping to the genes for samples with and without STVs.
- Of these genes, 538 showed a significant difference in gene expression (*q value* < 0.1).
- In the 538 genes, genes with terms of 'cell cycle' and 'mitotic cell cycle' were significantly enriched.




Brown; gene with *q-value* < 0.1

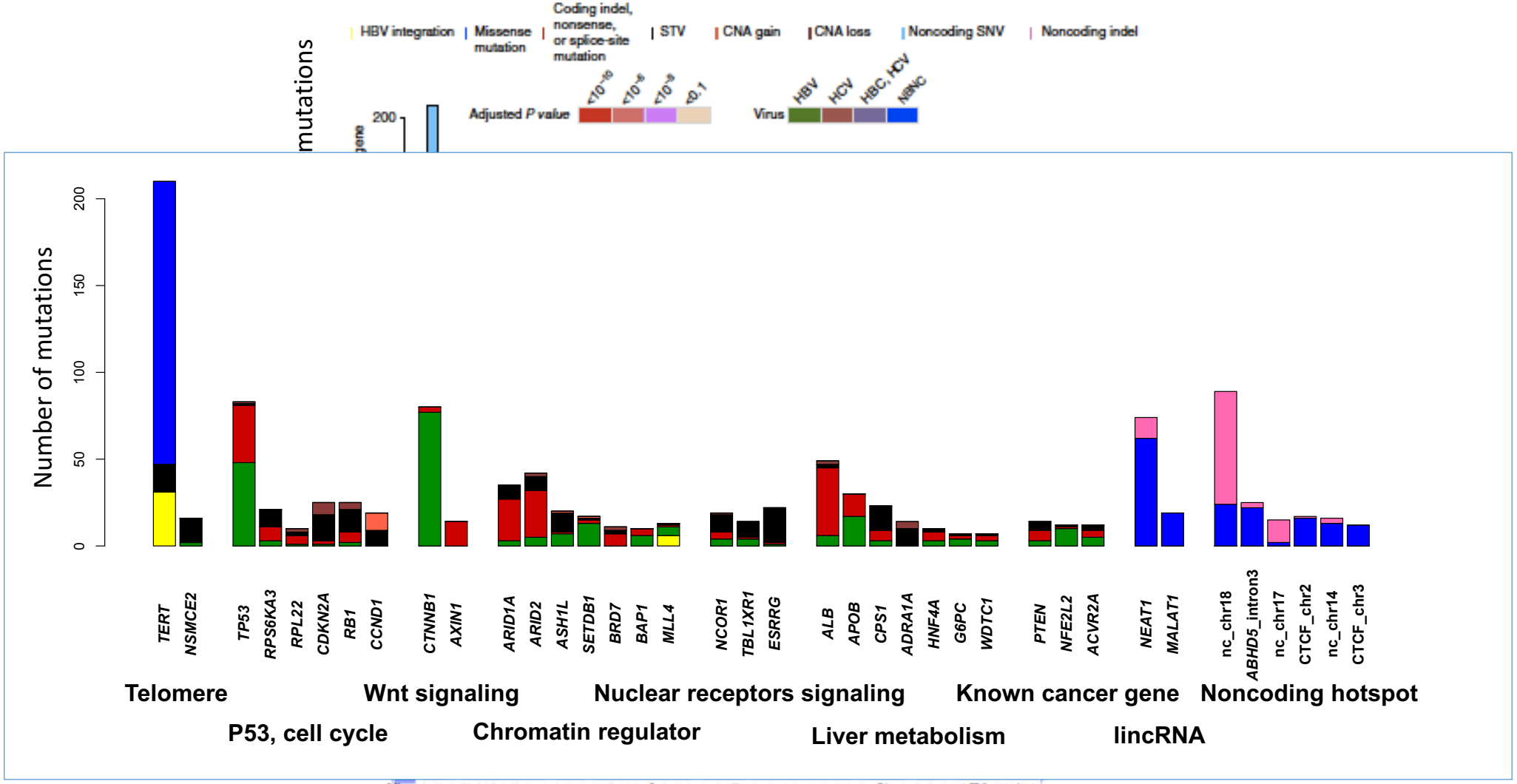
Statistical analysis was done with EdgeR package.

# Mutation in *TERT* gene and expression




- Promoter mutation (n=163), STVs (n=16), and HBV integrations (n=14) were identified in the *TERT* gene body and promoter region .
- Type of STVs were translocation and inversion.
- STVs, HBV and promoter point mutations were mutually exclusive.

# Summary of mutations in liver cancer



# Summary

- We analyzed whole-genome landscape of somatic alterations in 300 liver cancers.
- The number of point mutations were ~3/Mbp. But one sample (RK308) showed a hypermutated phenotype due to mismatch-repair deficiency.
- Our analysis identified 38 significantly mutated genes. Of these, chromatin regulators (*ASH1L,* and *SETDB1*), nuclear receptors (*TBL1XR1, NCOR1*, and *ESRRG*), DNA mismatch repair genes (*NSMCE2,* and *MACROD2*) and genes related to liver metabolism (*HNF4A,* and *G6PC*) were observed.
- Six long intergenic noncoding RNA (lincRNA) genes (*NEAT1* and *MALAT1*), ten promoters (*TFPI2*, *MED16*, and *WDR74*) and nine UTRs (*BCL6* and *AFF4*) were identified as regions with a significantly large number of mutations.
- A strong association between gene expression and STVs was observed.
- In the *TERT* regions, recurrent STVs and HBV integrations were detected, and associated with gene expression.

YOKOZUNA
(Grand Champ)
横綱

TP53, CTNNB1, TERT
(SNV, STV, HBV integration, amplification)

OH-ZEKI
(Champ)
大関

Wnt pathway genes
AXIN1, APC…
HBV integration
lincRNA?

SEKI-WAKE
関脇

Chromatin Regulators
ARID2, ARID1A, MLL, MLL4,…

Many long-tail genes
( 幕内、十両）

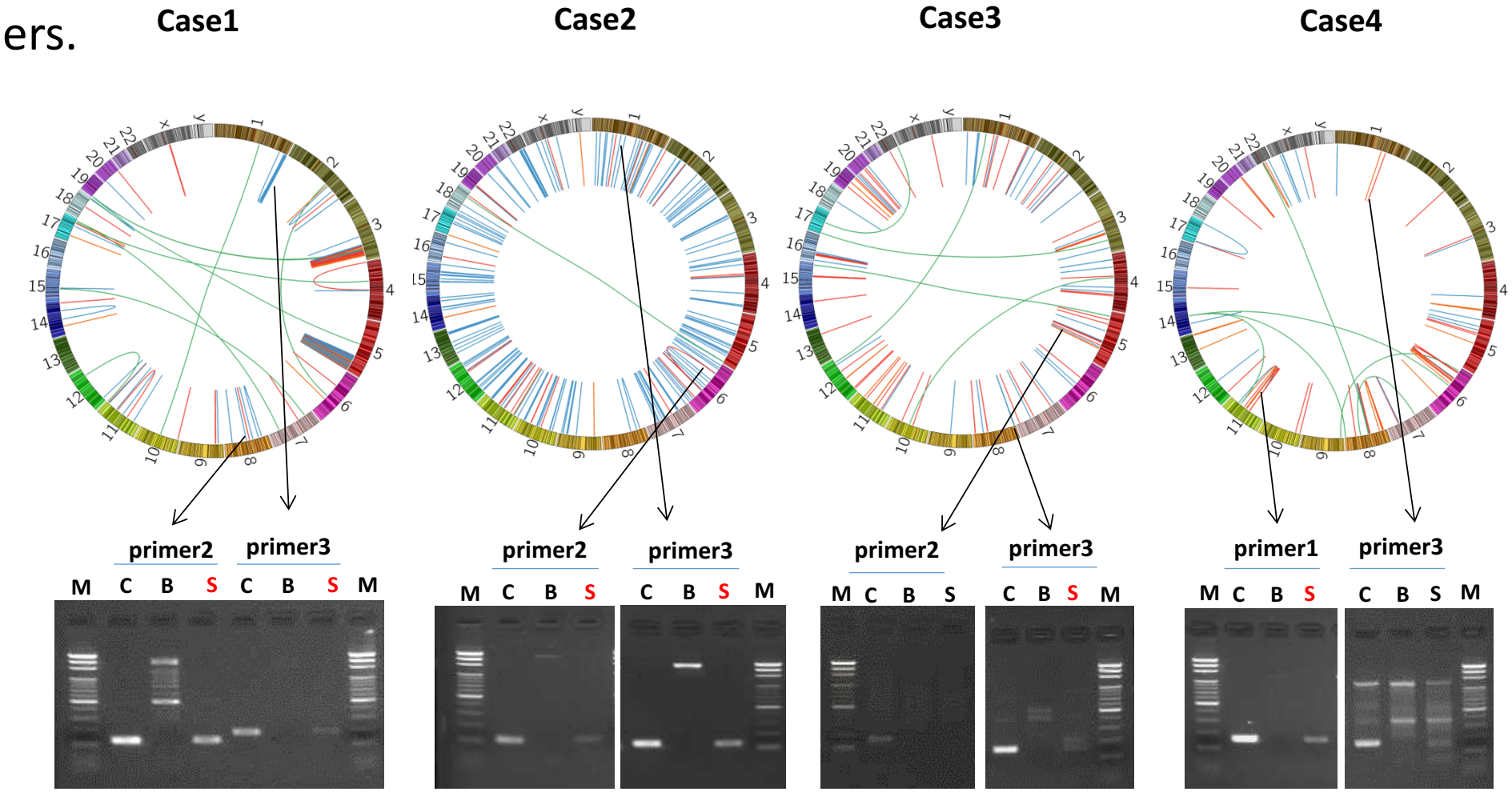
# Circulating tumor DNA analysis for liver cancers

## Collaboration with Hiroshima University

Ono et al. **Cellular and Molecular Gastroenterology and Hepatology (2015)**

# “Personal” blood biomarkers to monitor the disease (1)

- Circulating tumor DNA (ctDNA) can be a good “personal” blood biomarker to monitor cancer.
- We detected ctDNA by structural variation.
- From WGS of 46 primary tumors, we selected STVs with high clonal proportion and designed PCR primers.



# “Personal” blood biomarkers to monitor the disease (2)

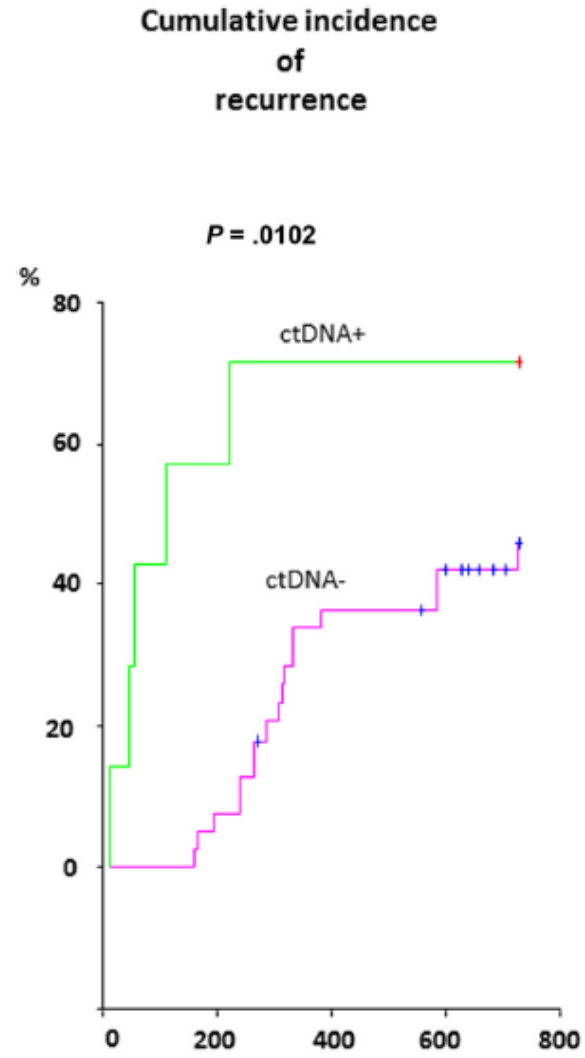
- In the cfDNA from 46 patients, STV was detected by PCR in seven patients before surgery.
- The cumulative incidence of recurrence in the ctDNA-positive group was significantly worse than that in the ctDNA-negative group.

**Table 2.** Summary of Clinical Characteristics of Patients

| Characteristic | Status of ctDNA Before Surgery | ctDNA Positive | ctDNA Negative | *P* Value |
|---|---|---|---|---|
| No. of patients | 46 | 7 | 39 | |
| Age at diagnosis (y) | 67 (32–89) | 68 (51–86) | 67 (32–89) | .426 |
| Gender (M/F) | 35/11 | 5/2 | 30/9 | .541[a] |
| Etiology (HBV/HCV(SVR)/NBNC) | 11/25 (6)/10 | 1/4(2)/2 | 10/21 (4)/8 | .775[b] |
| AFP (ng/mL) before surgery | 60.5 (<0.5–57,410) | 10,100 (12.7–57,410) | 15.8 (<5–35330) | .004[c] |
| DCP (mAU/mL) before surgery | 57.5 (2.6–135,640) | 23,156 (866–135,640) | 37 (2.6–16123) | <.001[c] |
| Tumor size (mm) | 25.5 (10–140) | 73 (35–140) | 23 (10–125) | <.001[c] |
| VV (0/1) | 38/8 | 4/3 | 34/5 | .089[a] |
| VA (0/1) | 45/1 | 6/1 | 39/0 | .152[a] |

Note: AFP, α-fetoprotein; DCP, des-γ-carboxy prothrombin; HBV, hepatitis B virus; HCV (SVR), hepatitis C virus sustained virologic response; NBNC, neither HBsAg (+) nor anti-HCV (+); VA, hepatic artery; VP, microscopic vascular invasion to portal vein; VV, hepatic vein.
[a]Fisher exact test.
[b]Pearson chi-square test.
[c]Mann-Whitney *U* test.

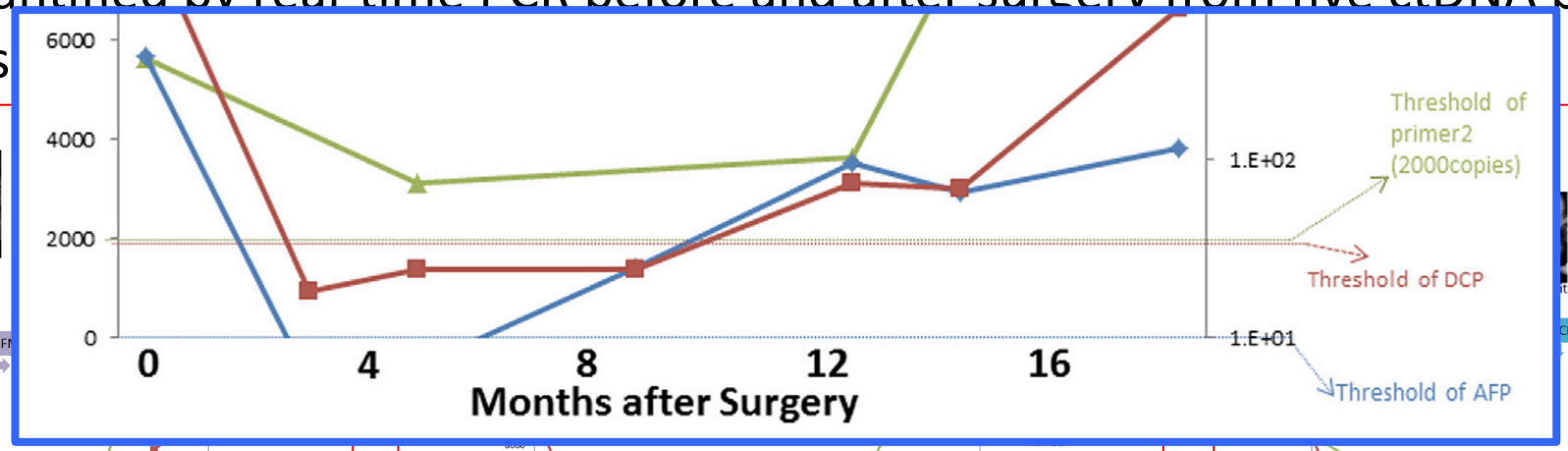
- **AFP and DCP levels and the rate of VP were significantly different between the cfDNA positive and negative samples.**

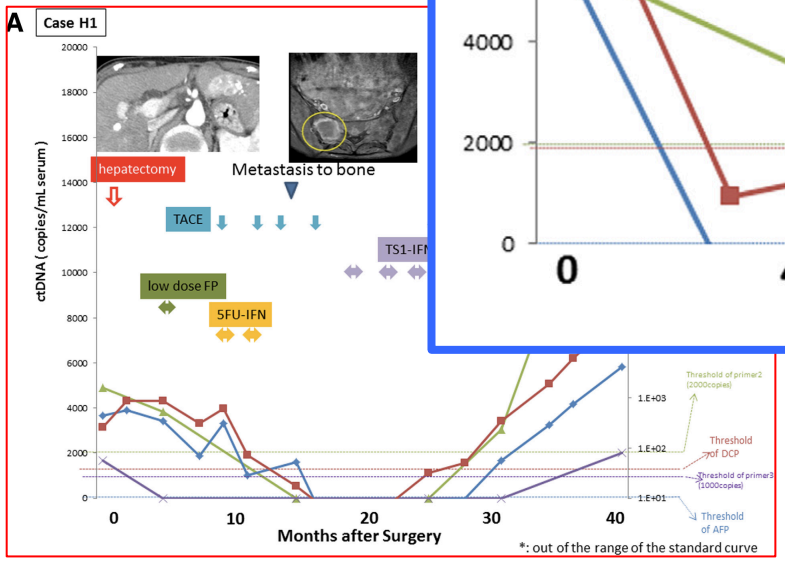



Ono et al. **Cellular and Molecular Gastroenterology and Hepatology (2015)**

# Quantification of ctDNA in serum

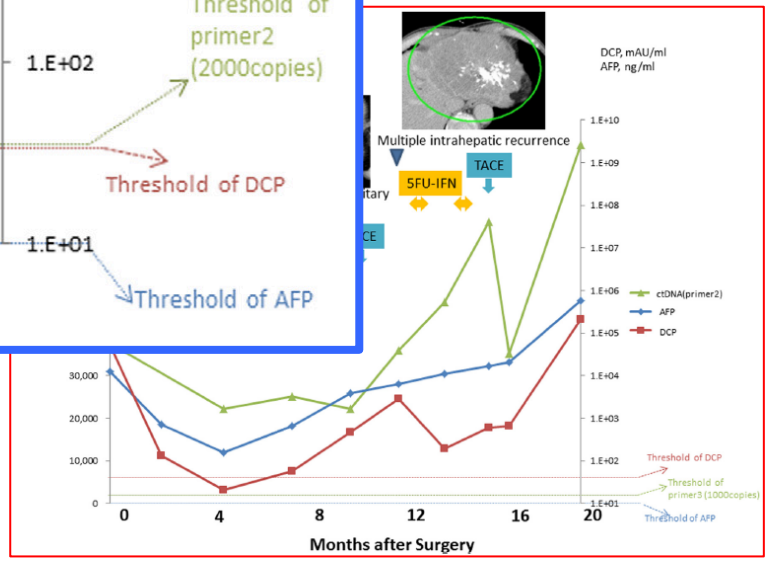
- ctDNA was quantified by real-time PCR before and after surgery from five ctDNA positive samples.
- The time cours

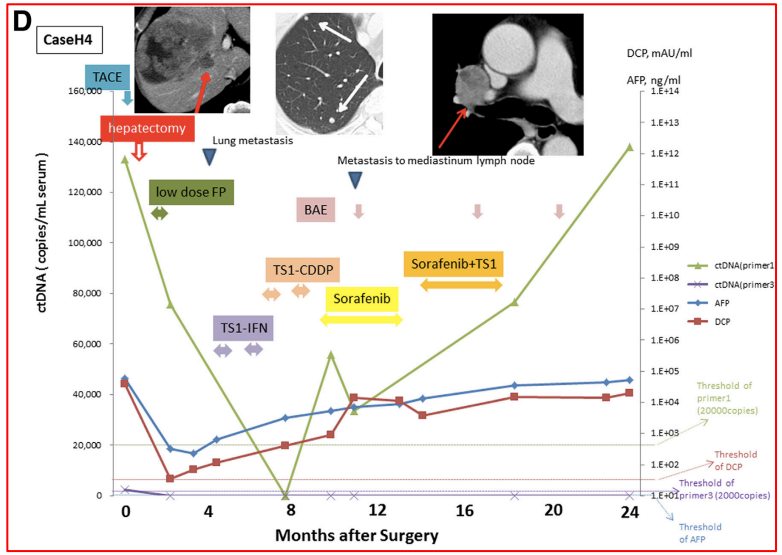


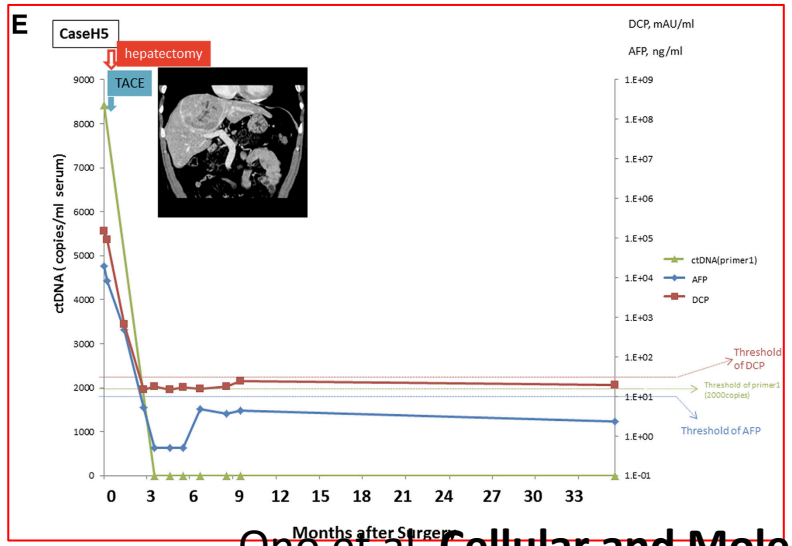










- In a patient, AFP and DCP had been negative between 3 and 5 months after the surgery, but ctDNA was detected. ctDNA may be more sensitive than the conventional biomarkers in some patients.

Ono et al. **Cellular and Molecular Gastroenterology and Hepatology (2015)**

# Collaborators

- **RIKEN**

**Mayuko Furuta**, **Tatsuhiko Tsunoda**, Hiroaki Taniguchi, Christopher P Wardell, Keith A Boroevich, Tetsuo Abe, Kaoru Nakano, Kazuhiro Maejima, Aya Sasaki-Oku, Ayako Ohsawa, Michiaki Kubo, **Hidewaki Nakagawa**

- **NCC**

**Yasushi Totoki**, **Mamoru Kato**, Hiromi Nakamura, Natsuko Hama, Fumie Hosoda, Yasuhito Arai, Shoko Ohashi, Nobuyoshi Hiraoka, Takuji Okusaka, Kazuaki Shimada, **Tatsuhiro Shibata**

- **University of Tokyo**

**Yuichi Shiraishi**, Hiroko Tanaka, Tetsuo Shibuya, Tomoko Urushidate, Genta Nagae, Shogo Yamamoto, Hiroki Ueda, Kenji Tatsuno, **Satoru Miyano**, **Hiroyuki Aburatani**

- **Hiroshima University School of Medicine**

Atsushi Ono, Yoshiiku Kawakami, Hiroshi Aikata, Koji Arihiro, Hideki Ohdan, Kazuaki Chayama

- **Hokkaido University Graduate School of Medicine**

Toru Nakamura, Satoshi Hirano

- **Keio University School of Medicine**

Hidenori Ojima

- **Wakayama Medical University**

Masaki Ueno, Shinya Hayami, Hiroki Yamaue

- **Osaka Medical Center for Cancer and Cardiovascular Diseases**

Kunihito Gotoh, Shigeru Marubashi, Shigeru Marubashi, Osamu Ishikawa

- **Tokyo Women's Medical University**

Shun-ichi Ariizumi, Masakazu Yamamoto