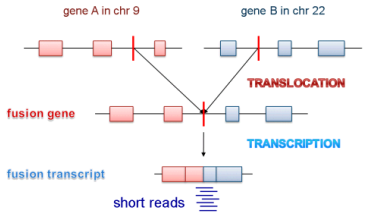
# 前立腺癌での融合遺伝子からの遺伝子発現

次世代シークエンサーによる癌部／非癌部のRNA塩基配列解析

# 2. 融合遺伝子転写産物(Fusion gene transcripts)




https://en.wikipedia.org/wiki/Fusion_gene

History

The first fusion gene[1] was described in cancer cells in the early 1980s. The finding was based on the discovery in 1960 by Peter Nowell and David Hungerford in Philadelphia of a small abnormal marker chromosome in patients with chronic myeloid leukemia - the first consistent chromosome abnormality detected in a human malignancy, later designated the Philadelphia chromosome.[2] In 1973, Janet Rowley in Chicago showed that the Philadelphia chromosome had originated through a translocation between chromosomes 9 and 22, and not through a simple deletion of chromosome 22 as was previously thought. Several investigators in the early 1980s showed that the Philadelphia chromosome translocation led to the formation of a new BCR/ABL1 fusion gene, composed of the 3'part of the ABL1 gene in the breakpoint on chromosome 9 and the 5' part of a gene called BCR in the breakpoint in chromosome 22. In 1985 it was clearly established that the fusion gene on chromosome 22 produced an abnormal chimeric BCR/ABL1 protein with the capacity to induce chronic myeloid leukemia.

At present, scientists have identified 358 gene fusions involving 337 different genes. These genes have been found in practically all main subtypes of human neoplasia.[3] The identification of these fusion genes play a prominent role in being a diagnostic and prognostic marker.[4]

Oncogenes

It has been known for 30 years that the corresponding gene fusion plays an important role in tumorigenesis.[5] Fusion genes can contribute to tumor formation because fusion genes can produce much more active abnormal protein than non-fusion genes. Often, fusion genes are oncogenes that cause cancer; these include BCR-ABL,[6] TEL-AML1 (ALL with t(12 ; 21)), AML1-ETO (M2 AML with t(8 ; 21)), and TMPRSS2-ERG with an interstitial deletion on chromosome 21, often occurring in prostate cancer.[7] In the case of TMPRSS2-ERG, by disrupting androgen receptor (AR) signaling and inhibiting AR expression by oncogenic ETS transcription factor, the fusion product regulate the prostate cancer.[8] Most fusion genes are found from hematological cancers, sarcomas, and prostate cancer.[9][10] BCAM-AKT2 is a fusion gene that is specific and unique to high-grade serous ovarian cancer.[11]

Oncogenic fusion genes may lead to a gene product with a new or different function from the two fusion partners. Alternatively, a proto-oncogene is fused to a strong promoter, and thereby the oncogenic function is set to function by an upregulation caused by the strong promoter of the upstream fusion partner. The latter is common in lymphomas, where oncogenes are juxtaposed to the promoters of the immunoglobulin genes.[12] Oncogenic fusion transcriptsmay also be caused by trans-splicing or read-through events.[13]

Since chromosomal translocations play such a significant role in neoplasia, a specialized database of chromosomal aberrations and gene fusions in cancer has been created. This database is called Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer[14]

# 融合遺伝子の転写産物の検出のためのプログラム: 現在39個あります

1. Barnacle: http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-14-550
2. Bellerophontes: http://bioinformatics.oxfordjournals.org/content/28/16/2114.long
3. BreakDancer: http://www.nature.com/nmeth/journal/v6/n9/abs/nmeth.1363.html
4. BreakFusion: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3389765/
5. BreakPointer: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3561864/
6. ChimeraScan: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3187648/
7. Comrad: http://bioinformatics.oxfordjournals.org/content/27/11/1481.long
8. CRAC: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4053775/
9. deFuse: http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1001138
10. Dissect: http://bioinformatics.oxfordjournals.org/content/28/12/i179.abstract
11. EBARDenovo: http://bioinformatics.oxfordjournals.org/content/early/2013/03/01/bioinformatics.btt092
12. EricScript: http://bioinformatics.oxfordjournals.org/content/28/24/3232
13. FusionAnalyser: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3439881/
14. FusionCatcher: http://biorxiv.org/content/early/2014/11/19/011650.full-text.pdf+html
15. FusionFinder: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3384600/
16. FusionHunter: http://bioinformatics.oxfordjournals.org/content/27/12/1708.long
17. FusionMap: http://bioinformatics.oxfordjournals.org/content/27/14/1922
18. FusionQ: http://www.biomedcentral.com/1471-2105/14/193
19. FusionSeq: http://www.genomebiology.com/2010/11/10/R104
20. IDP-fusion: http://nar.oxfordjournals.org/content/early/2015/06/03/nar.gkv562.full

..

..

39. ViralFusionSeq: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3582262/

# 融合遺伝子の転写産物の検出のためのプログラム: どれが良いのか

Fusion gene検出方法の性能評価 日本人類遺伝学会第57回大会(2012年10月) ポスター http://www.rikengenesis.jp/ori/50279/pdf/poster_1.pdfより

一般的に、Fusion gene同定ソフトウェアのアルゴリズムは、ペアエンド・リードを使うType1とシングルエンド・リードを使うType2 の2種類に分かれる。Type1はペアのリードが異なる遺伝子上にマッピングされた現象からFusion geneを推定する。一方、Type2はFusion junctionをまたいでいるリードからFusion geneを推定する。ペアエンド・リードを利用するソフトウェアは、Type1とType2 の両方に対応可能である。

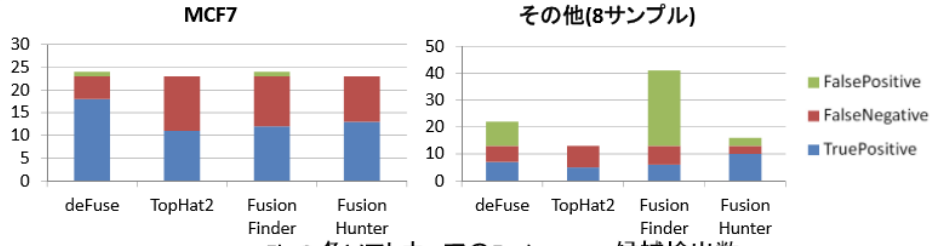


Fig.2 各ソフトウェアのFusion gene候補検出数

Table.4 既知のFusion gene検出結果

MCF7

| Sample | Gene name 1 | Gene name 2 | deFuse | TopHat 2 | Fusion Finder | Fusion Hunter |
|---|---|---|---|---|---|---|
| MCF7 | DEPDC1B | ELOVL7 | ○ | ○ | ○ | ○ |
| MCF7 | RPS6KB1 | TMEM49 | ○ | ○ | ○ | ○ |
| MCF7 | GCN1L1 | MSI1 | ○ | ○ | ○ | ○ |
| MCF7 | ESR1 | C6orf97 | ○ | ○ | ○ | ○ |
| MCF7 | EIF3D | MYH9 | ○ | ○ | ○ | ○ |
| MCF7 | ABCC1 | C16orf45 | ○ | ○ | ○ | ○ |
| MCF7 | SULF2 | ARFGEF2 | ○ | ○ | ○ | ○ |
| MCF7 | PRICKLE2 | SULF2 | ○ | ○ | ○ | ○ |
| MCF7 | RPS6KB1 | DIAPH3 | ○ | ○ | ○ | × |
| MCF7 | ADAMTS19 | SLC27A6 | ○ | × | ○ | ○ |
| MCF7 | BCAS4 | BCAS3 | △ | ○ | ○ | ○ |
| MCF7 | TAF4 | BRIP1 | ○ | × | ○ | ○ |
| MCF7 | NUP210L | GATAD2B | ○ | ○ | × | ○ |
| MCF7 | SLC25A24 | NBPF6 | ○ | × | ○ | × |
| MCF7 | FAM171A2 | ATXN7L3 | ○ | × | × | ○ |
| MCF7 | ABCA5 | PPP4R1L | ○ | × | × | ○ |
| MCF7 | SYTL2 | PICALM | ○ | × | × | × |
| MCF7 | TBL1XR1 | RGS17 | ○ | × | × | × |
| MCF7 | AHCYL1 | RAD51C | ○ | × | × | × |
| MCF7 | USP31 | CRYL1 | × | × | × | × |
| MCF7 | C16orf62 | IQCK | × | × | × | × |
| MCF7 | POP1 | MATN2 | × | × | × | × |
| MCF7 | MYO6 | SENP6 | × | × | × | × |
| Number of Detected Fusion genes | | | 18 | 11 | 13 | 14 |

その他(8サンプル)

| Sample | Gene name 1 | Gene name 2 | deFuse | TopHat 2 | Fusion Finder | Fusion Hunter |
|---|---|---|---|---|---|---|
| NCIH660 | TMPRSS2 | ERG | ○ | ○ | ○ | ○ |
| M000216 | KCTD2 | ARHGEF12 | × | × | × | ○ |
| M990802 | ANKHD1 | CYSTM1 | △ | × | ○ | ○ |
| M980409 | PLA2G1B | GCN1L1 | × | × | × | ○ |
| M010403 | WDR72 | SCAMP2 | × | × | × | × |
| 501Mel | SLC12A7 | C11orf67 | ○ | ○ | ○ | ○ |
| 501Mel | PARP1 | MIXL1 | × | ○ | × | ○ |
| 501Mel | GNA12 | SHANK2 | ○ | × | × | ○ |
| M000921 | RECK | ALX3 | ○ | × | × | ○ |
| M000921 | TLN1 | C9orf127 | × | × | × | × |
| K562 | BCR | ABL1 | ○ | ○ | ○ | ○ |
| K562 | NUP214 | XKR3 | ○ | ○ | ○ | ○ |
| K562 | BAG6 | SLC44A4 | ○ | × | ○ | × |
| Number of Detected Fusion genes | | | 7 | 5 | 6 | 10 |

○:検出、△:Filterにより除去、×:非検出

## 3-3. 結果:ソフトウェアの組合せによる検出率の向上

既知のFusion geneについて、2つのソフトウェアの組み合わせによる検出数は以下のようになった(Fig.4)。deFuseとFusionHunterの組み合わせが最も検出率が高い。

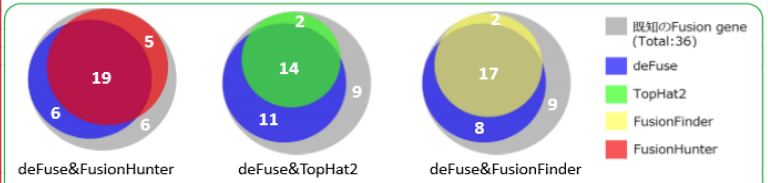
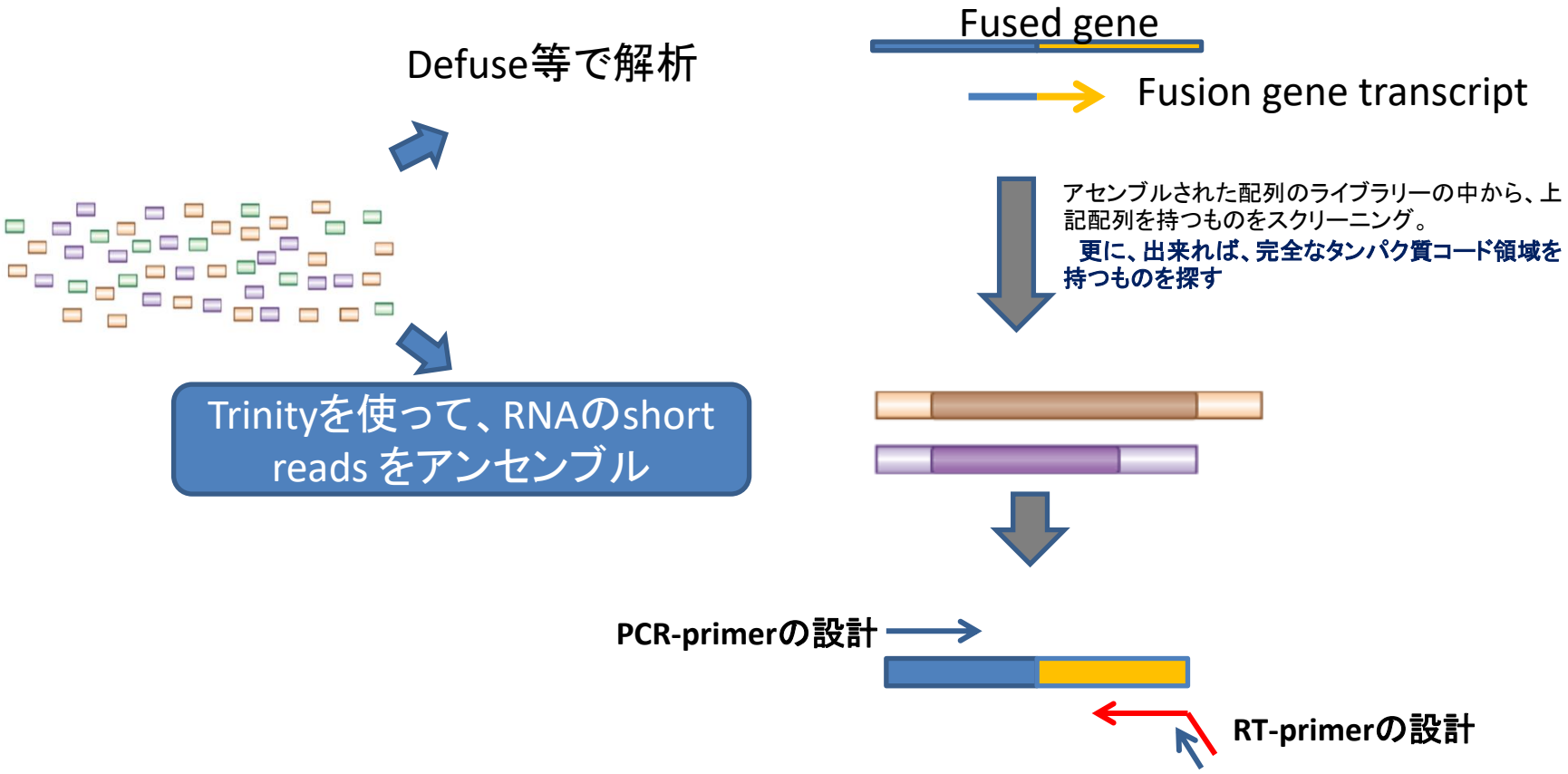


Fig.4 2つのソフトウェアの組み合わせによる検出数

・ペアエンドに対応したソフトウェアを推奨

Type1とType2の両方のアルゴリズムに対応したdeFuseとFusionHunterは、Type1のみに対応しているTopHat2やFusionFinderよりも、より多くのFusion geneの同定に成功している。この傾向は、転写量の低い遺伝子や、データ量の少ないデータで、顕著になると思われる。

# Fusion geneの検証

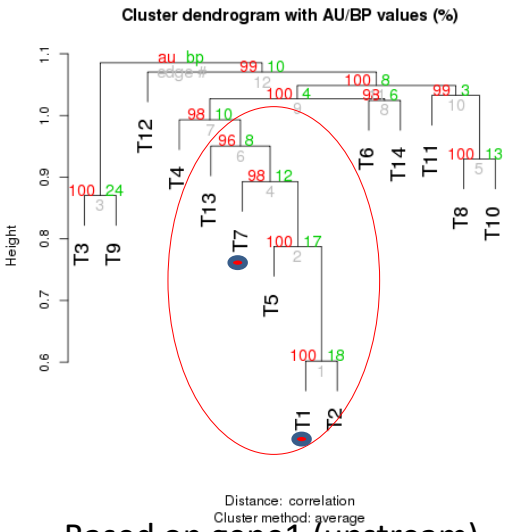


# 癌(14サンプル)における融合遺伝子出現頻度

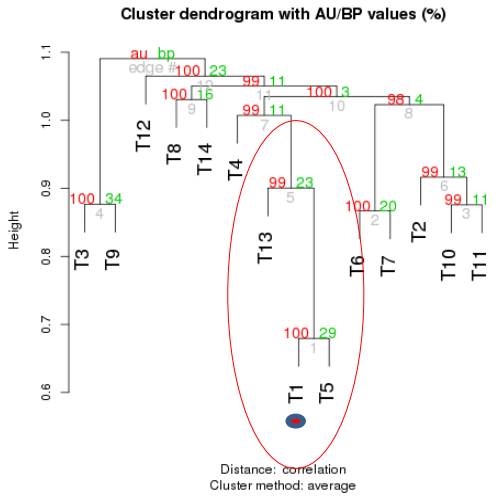
**前立腺細胞の癌化の過程で、染色体の再編により、正常細胞では存在しない、新たな遺伝子の融合が起こり、融合遺伝子の転写産物が生じる。これらが細胞の癌化に関連している可能性がある。また、 この転写産物を利用して、転移癌か原発癌かの解析を行うことが可能**

| gene1 | gene1type | gene1name | gene2 | gene2type | gene2name | normal | tumor | |
|---|---|---|---|---|---|---|---|---|
| ENSG00000160862 | coding | AZGP1 | ENSG00000176402 | coding | GJC3 | 0 | 1 | patient3 |
| ENSG00000160862 | coding | AZGP1 | ENSG00000176402 | coding | GJC3 | 0 | 1 | patient7 |
| ENSG00000160862 | coding | AZGP1 | ENSG00000176402 | intron | GJC3 | 0 | 1 | patient8 |
| ENSG00000160862 | coding | AZGP1 | ENSG00000176402 | intron | GJC3 | 0 | 1 | patient9 |
| ENSG00000160862 | coding | AZGP1 | ENSG00000176402 | coding | GJC3 | 0 | 1 | patient11 |
| ENSG00000160862 | coding | AZGP1 | ENSG00000176402 | intron | GJC3 | 0 | 1 | patient12 |
| ENSG00000160862 | coding | AZGP1 | ENSG00000176402 | intron | GJC3 | 0 | 1 | patient14 |
| ENSG00000198131 | coding | ZNF544 | ENSG00000083842 | upstream | ZNF8 | 0 | 1 | patient5 |
| ENSG00000198131 | coding | ZNF544 | ENSG00000083842 | upstream | ZNF8 | 0 | 1 | patient6 |
| ENSG00000198131 | coding | ZNF544 | ENSG00000083842 | upstream | ZNF8 | 0 | 1 | patient11 |
| ENSG00000198131 | coding | ZNF544 | ENSG00000083842 | upstream | ZNF8 | 0 | 1 | patient12 |
| ENSG00000198131 | coding | ZNF544 | ENSG00000083842 | upstream | ZNF8 | 0 | 1 | patient13 |
| ENSG00000198131 | coding | ZNF544 | ENSG00000083842 | upstream | ZNF8 | 0 | 1 | patient14 |
| ENSG00000198466 | intron | ZNF587 | ENSG00000204514 | coding | ZNF814 | 0 | 1 | patient3 |
| ENSG00000198466 | coding | ZNF587 | ENSG00000178935 | downstream | ZNF552 | 0 | 1 | patient5 |
| ENSG00000198466 | coding | ZNF587 | ENSG00000178935 | downstream | ZNF552 | 0 | 1 | patient6 |
| ENSG00000198466 | coding | ZNF587 | ENSG00000178935 | downstream | ZNF552 | 0 | 1 | patient7 |
| ENSG00000198466 | intron | ZNF587 | ENSG00000204514 | coding | ZNF814 | 0 | 1 | patient9 |
| ENSG00000198466 | coding | ZNF587 | ENSG00000178935 | downstream | ZNF552 | 0 | 1 | patient11 |
| ENSG00000198952 | coding | SMG5 | ENSG00000163472 | intron | TMEM79 | 0 | 1 | patient3 |
| ENSG00000198952 | utr3p | SMG5 | ENSG00000160781 | utr5p | PAQR6 | 0 | 1 | patient3 |
| ENSG00000198952 | utr3p | SMG5 | ENSG00000160781 | utr5p | PAQR6 | 0 | 1 | patient7 |
| ENSG00000198952 | utr3p | SMG5 | ENSG00000160781 | utr5p | PAQR6 | 0 | 1 | patient10 |
| ENSG00000198952 | utr3p | SMG5 | ENSG00000160781 | utr5p | PAQR6 | 0 | 1 | patient11 |
| ENSG00000184012 | coding | TMPRSS2 | ENSG00000157554 | coding | ERG | 0 | 1 | patient1 |
| ENSG00000184012 | downstream | TMPRSS2 | ENSG00000130177 | intron | CDC16 | 0 | 1 | patient2 |
| ENSG00000184012 | coding | TMPRSS2 | ENSG00000157554 | coding | ERG | 0 | 1 | patient5 |
| ENSG00000184012 | coding | TMPRSS2 | ENSG00000123358 | intron | NR4A1 | 0 | 1 | patient7 |
| ENSG00000184012 | coding | TMPRSS2 | ENSG00000157554 | coding | ERG | 0 | 1 | patient13 |
| ENSG00000127863 | intron | TNFRSF19 | ENSG00000027001 | downstream | MIPEP | 0 | 1 | patient3 |
| ENSG00000127863 | intron | TNFRSF19 | ENSG00000027001 | downstream | MIPEP | 0 | 1 | patient10 |
| ENSG00000127863 | intron | TNFRSF19 | ENSG00000027001 | downstream | MIPEP | 0 | 1 | patient12 |
| ENSG00000127863 | intron | TNFRSF19 | ENSG00000027001 | downstream | MIPEP | 0 | 1 | patient13 |
| ENSG00000127863 | intron | TNFRSF19 | ENSG00000027001 | downstream | MIPEP | 0 | 1 | patient14 |
| ENSG00000140006 | upstream | WDR89 | ENSG00000200693 | intron | U3 | 0 | 1 | patient3 |
| ENSG00000140006 | upstream | WDR89 | ENSG00000200693 | intron | U3 | 0 | 1 | patient5 |
| ENSG00000140006 | upstream | WDR89 | ENSG00000200693 | intron | U3 | 0 | 1 | patient6 |
| ENSG00000140006 | upstream | WDR89 | ENSG00000200693 | intron | U3 | 0 | 1 | patient7 |
| ENSG00000140006 | upstream | WDR89 | ENSG00000200693 | intron | U3 | 0 | 1 | patient13 |
| ENSG00000172175 | intron | MALT1 | ENSG00000267501 | upstream | RP11-108P20.2 | 0 | 1 | patient3 |
| ENSG00000172175 | intron | MALT1 | ENSG00000267501 | upstream | RP11-108P20.2 | 0 | 1 | patient8 |
| ENSG00000172175 | intron | MALT1 | ENSG00000267501 | upstream | RP11-108P20.2 | 0 | 1 | patient11 |
| ENSG00000172175 | intron | MALT1 | ENSG00000267501 | upstream | RP11-108P20.2 | 0 | 1 | patient12 |
| ENSG00000172175 | intron | MALT1 | ENSG00000267501 | upstream | RP11-108P20.2 | 0 | 1 | patient14 |

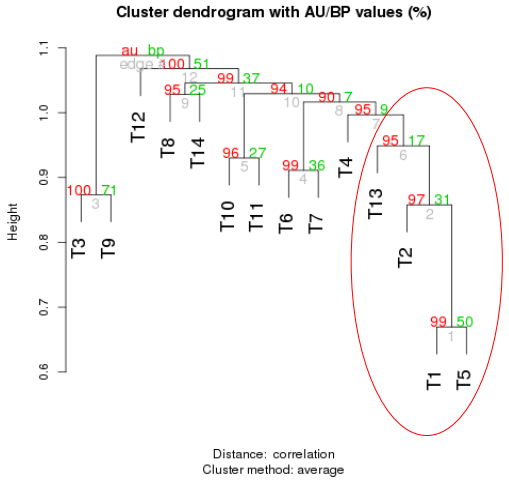
# Fusion gene による癌のクラスタリング



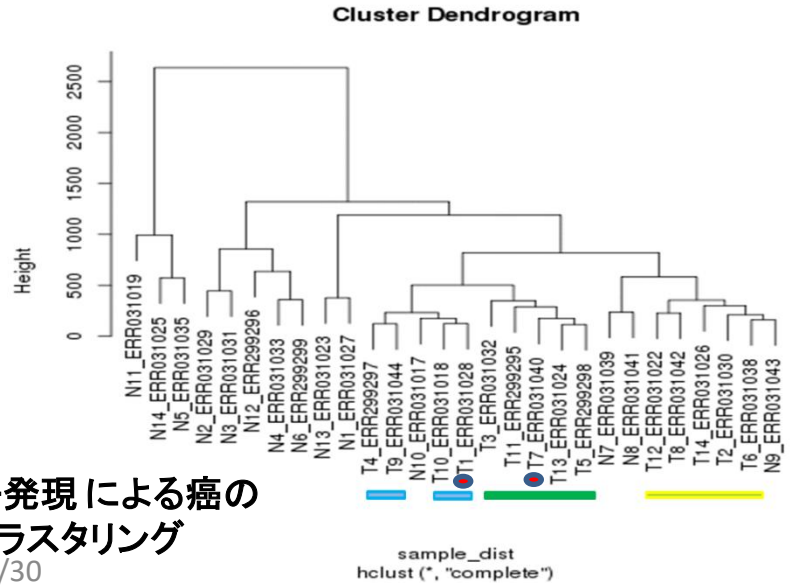

Based on gene1 (upstream)




Based on gene2 (downstream)




Based on genes 1 and 2




**遺伝子発現による癌のクラスタリング**

Supplementary Patient information

| No. | Age | Preoperative PSA | Stage | Gleason Score | Metastasis |
|---|---|---|---|---|---|
| 1 | 74 | 9.85 | T2cN0M0 | 3+4 | 0 |
| 2 | 73 | 1.36 | T1cN0M0 | 3+3 | 0 |
| 3 | 71 | 9.62 | T2aN0M0 | 2+2 | 0 |
| 4 | 54 | 7.44 | T2cN0M0 | 3+3 | 0 |
| 5 | 62 | 7.76 | T4N0M0 | 3+4 | Bladder |
| 6 | 69 | 4.04 | T1cN0M0 | 3+4 | 0 |
| 7 | 52 | 30.33 | T3bN0M0 | 3+4 | 0 |
| 8 | 66 | 10.4 | T3aN0M0 | 3+4 | 0 |
| 9 | 56 | 9.78 | T2cN0M0 | 3+4 | 0 |
| 10 | 75 | 10.93 | T2cN0M0 | 3+4 | 0 |
| 11 | 57 | 6.99 | T2cN0M0 | 3+4 | 0 |
| 12 | 80 | 22.38 | T1cN0M0 | 4+3 | 0 |
| 13 | 75 | 12.69 | T4N0M0 | 5+3 | Bladder |
| 14 | 73 | 12.8 | T2bN0M0 | 3+2 | 0 |

- 転移の可能性?