

博士論文

次世代シーケンサーを用いた

レトロウィルス **Clonality** 解析システムの構築

安原（青木） 桜

目次

背景.....	3
実験材料と方法	8
実験結果	17
考察.....	24
<謝辞>	
<引用文献>	

背景

HTLV-1 と ATL

HTLV-1 はヒトにおいて最初に同定されたレトロウイルスであり、成人 T 細胞白血病 (Adult T-cell leukemia ATL) の原因ウイルスである。その他に、HAM (HTLV-1 associated myelopathy)、TSP (tropical spastic paraparesis) と呼ばれる脊髄痙性麻痺症や目の炎症であるぶどう膜炎 (HTLV-1 uveitis, HU) の原因であることも明らかとなった。ATL は HTLV-1 感染 T リンパ球が腫瘍化することで発症する白血病/リンパ腫であり、未だ有効な治療法が無く、予後が悪いことで知られている (1~3)。

HTLV-1 の感染者は日本では九州、沖縄を中心とした西南部に多く分布している。また近年では人の移動が増加したため都市部にも感染者が多く分布している。世界的にはカリブ海沿岸、南アメリカ、中央アフリカに多くの感染者が存在する(4,5)。日本は 108 万人の HTLV-1 既感染者が存在するといわれている。その内数% (2~5%) が約 50 年以上の長い潜伏期間を経て ATL を発症するといわれており、年間約 1000 人以上が ATL を発症している(6)。

ATL は様々な病態、臨床経過をとり、①急性型、②リンパ腫型、③慢性型、④くすぶり型の 4 つの病型と急性転化の 1 つの病態に分類される(7,8)。ATL の病型は急性型、慢性型は早急な治療を必要とする Aggressive タイプと経過観察を行う慢性型、くすぶり型の Indolent タイプに大きく分類することができる。Indolent タイプの ATL 患者は症状を伴わない場合が多いが明らかに無症候性キャリアに比べて予後が悪く、急性転化する場合も多いため十分な経過観察が必要である(9-11)。ATL の特徴的な症状に腫瘍細胞の臓器浸潤や高カルシウム血症がある(12)。

ATL と多段階発がんモデル

悪性腫瘍の成り立ちは一般的に複数の過程を経て(multi-step)、年余をかけて進んでいくものと考えられる。

ATL は原因によって明らかに定義される唯一の腫瘍であるため、HTLV-1 感染により起こる ATL は多段階発癌過程を解析するよいモデルであると考えられる。実際疫学的解析により、ATL 発症は 5 つ以上の genetic event の集積による典型的な多段階発癌モデルに合致することが示されている (Fig1)。

HTLV-1 は乳幼児期に母乳を介して感染し、ポリクローナルな増殖を繰り返し

た後感染細胞ごとにクローンを構成する。長い年月の中で複数の genetic event を経た後 1 つのクローンがモノクローナルな増殖を示し ATL を発症すると言われている(13)。

ATL クローンの特徴

ATL 細胞はサザンブロット法で解析すると、HTLV-1 のモノクローナルな組み込みが検出される(14-16)。従って、HTLV-1 感染細胞が腫瘍化してモノクローナルに増殖したものと考えられている。サザンブロット法で複数のポリクローナルなプロウィルスバンドが検出される症例も認められるが、この場合でも TCR のリアレンジメントの解析や定量的 PCR による細胞あたりのプロウィルスコピー数の検討から、モノクローナルな ATL 細胞に複数のプロウィルスが組み込まれていることが多い(13)。しかし、症例報告などを調べると同一個体に同時あるいは継時的に複数の ATL クローンが検出される例がある。以下に例を挙げる。

- ①CD4+ CD8- と CD4- CD8+ の 2 つのクローンが認められた慢性型 ATL (15)、
- ②自然寛解後に異なったクローンにより再発した例(16)、
- ③5 個の異なったクローンが認められる巨大な皮膚腫留を呈した例(17)などである。

このような同一個体内に複数の ATL 細胞クローンが存在することの意義は明らかではないが、感染個体内で複数の感染細胞クローンがほぼ同時に腫瘍化の過程をたどっている事を示唆するデータと考えることが可能である。

ATL は多剤併用化学療法によって短期間の寛解が得られるのみで非常に短期間で再び ATL 細胞の増殖が認められる(18,19)。このことは ATL 細胞が非常に用意に薬剤耐性を獲得することによると考えられている。しかし、この現象も新たな ATL クローンの増殖を観察しているに過ぎないという可能性もあり、この点の検討はこれまで全くなされていない。

無症候性キャリアの HTLV-1 感染細胞クローンの特徴

無症候性キャリアの HTLV-1 感染細胞は生涯 2~3% であると言われている (Fig2)。それゆえ無症候性キャリアの場合モノクローナルな HTLV-1 感染クローンは存在せず、複数のクローンがポリクローナルに存在すると考えられている(20-24)。しかし無症候性キャリアの HTLV-1 感染細胞数は少ないため HTLV-1 感染クローンの検出が困難となりクローンの特徴解析はほとんどなされていない。

ウィルスロード

ヒトの T リンパ球に HTLV-1 が感染すると、1 細胞あたり 1 個のプロウイルスがランダムに組み込まれる(14,25,26)。HTLV-1 感染者の血液中の全リンパ球数とプロウイルス数を Real time PCR により測定することで全リンパ球中の HTLV-1 感染細胞の含有率が測定できる(27) (Fig3)。ほとんどの無症候性キャリアの場合、全リンパ球中約 1%前後が感染細胞である(28-30)。この感染細胞率(ウィルスロード)が 4%を超えるキャリアからのみ ATL を発症することが大規模コホート研究(JSPFAD)より明らかとなった(31)。しかし 4%を超える無償構成キャリアの割合は全キャリアの 1/4 であり、その多くは生涯 ATL を発症しない場合も多い。しかし、ウィルスロードを用いた ATL リスク評価系はこれ以上のキャリアの分類を行うことは不可能である。

クローナリティ解析

ATL クローン候補を早期に検出するためにはモノクローナルな増殖を示す HTLV-1 感染細胞を無症候性キャリアの検体から検出することが必要となる。上記の条件を満たすためには HTLV-1 感染細胞クローンを①網羅的に②特異的にそして、③各 HTLV-1 感染細胞クローンサイズを正確に計測するクローナリティ解析システムが必要である (Fig 4)。クローナリティとは各 HTLV-1 感染細胞クローンが全 HTLV-1 感染細胞に占める割合を意味する。

HTLV-1 感染クローンは大腸ポリープに例えることができる。大腸ポリープもほとんどの場合は良性で問題ないが、一部腫瘍化するリスクを持っている。大腸ポリープのような固形腫瘍の場合は視覚的に腫瘍の大きさを計測できるが、ATL は血液腫瘍のため HTLV-1 感染クローンサイズを視覚的に計測することはできない。

HTLV-1 はヒトゲノムに 1 コピー組み込まれる。各クローンは同一ヒトゲノム組み込み部位を 1 コピーもつ細胞 (Sister cell) で構成される(Fig 5)。つまり HTLV-1 組み込み部位を Sister cell 毎に分類することでクローンサイズを計測すること、つまりクローナリティ解析が可能となる。

既存クローナリティ解析システム

HTLV-1 感染細胞のクローナリティ解析はサザンブロット、もしくはレトロウィルスのヒトゲノム組み込み部位を増幅する **Inverse PCR 法**、**Nested splinkerette PCR 法**である(32,33)(Fig 6)。しかし、これらの方法では感度に限界があり、全 HTLV-1 キャリアのクローナリティ解析を行うことはできない。サザンブロット法はウィルスロードが数%以上の場合のみクローナリティ解析が可能であり、ウィルスロードの低い HTLV-1 キャリアのクローナリティ解析には利用できない。**Inverse PCR 法**の場合、PCR の増幅バンドの濃淡からクローナリティを判断することになり定量性に欠ける上、検出できるクローン数は数種類に留まることが多い。HTLV-1 キャリアのクローンは数十から数千と考えられており既存システムは一部のクローンを検出しているに過ぎない。また、**Inverse PCR** は制限酵素を用いるため制限酵素サイト近傍のクローンのみが検出される可能性が高い。

つまり既存のクローナリティ解析システムでは早期に **ATL** クローン候補を捉えることは不可能である。

次世代シーケンサーを用いた新規クローナリティ解析システム

2008 年ごろ登場した次世代シーケンサーは進化を続け利用範囲の裾野は広がり続けている。2009 年前後から **iPS 細胞** の **trans gene** のヒトゲノム組み込み部位の網羅的な解析が始まり (32,34)、続いて HTLV-1 と同じレトロウィルスである **HIV** のヒトゲノム組み込み部位の網羅的な解析が進んでいる(35,36)。次世代シーケンサーを用いることで網羅的に、かつ特異的にレトロウィルス感染細胞クローンの検出が可能となった。

2011 年 1 月に英国のグループより次世代シーケンサーを用いた HTLV-1 のクローナリティ解析システムが発表された(37)。HTLV-1 の組み込み部位を網羅的に次世代シーケンサーを用いてシーケンスすることでクローナリティ解析を実現している(Fig 7)。クローンを構成する **Sister cell** の計測にシーケンスサンプル調整時に生じる **DNA shear site** (断片化サイト) を利用している。各細胞を構成するゲノムを抽出し、**sonication** を用いて断片化する。ゲノム DNA はランダムに 1000bp 前後に断片化され、ゲノムの断片化サイトの末端配列は抽出した細胞毎に異なる。細胞毎に異なるゲノムの断片化サイトを解析することで **Sister cell** を計測している(Fig 8)。

新規クローナリティ解析システムの開発

上述のゲノム断片化サイトを用いるクローナリティ解析システムはクローンサイズの計測において問題点がある。それは断片化サイトが最大でも **400** サイト程度に留まるため、**400Sister cell** 以上で構成されるクローンのサイズ計測ができない点である (**Fig 9**)。英国のグループは統計処理にてこの問題点を考慮したクローナリティを導きだしているが、直接的な解析結果ではない。同程度の数の断片化サイトを持っているクローンが複数検出された場合も本当に同程度のクローンサイズであるとも限らない。

そこで我々はクローンサイズを正確に計測するクローナリティ解析システムを開発し、断片化部位を用いた解析と比較検討を試みた。また、急性型 **ATL**、くすぶり型 **ATL**、無症候性キャリアのクローナリティの違いについて検討した。

実験材料と方法

HTLV-1 プロウィルスゲノムの調整

医科研病院に入院していた患者 1 名について **Informed consent** を得た後、手術で摘出された固形腫瘍部から抽出したゲノム DNA を血液内科の内丸薫先生より分与された。JSPFAD コホート研究で集積されたサンプル DNA の抽出は以下のように行った。末梢血検体から **Ficoll-paque Plus(Amaersham Bioscience)**を用いた密度勾配遠心により、**PBMC** を分離した。分離した **PBM** から **QIAamp DNA Blood Mini Kit(QIAGEN)**を用いてゲノム DNA を抽出し、実験サンプルとして用いた。

Inverse PCR

ゲノム DNA 1 μ g を **Pst I**、**Xho I**、または **Eco RI** (いずれも **Takara**) を用いて切断し、反応物をエタノール沈殿後、**TE(10mM Tris-HCl Ph 8.0 EDTA) buffer** に溶解した。**T4 Ligase(Takara)**を用いて反応物全量を 4 $^{\circ}$ Cで一晩インキュベーションすることで切断したゲノム DNA を **Self-ligation** した。**Self-ligation** する際、切断したゲノム DNA 分子が隣接する他のゲノム DNA と **ligation** しないよう、反応液の容量を **total 250 μ l** とした。反応後、反応液をフェノール抽出し、エタノール沈殿後、**TE buffer** に溶解しその一部を鋳型として用いた。また **EcoRI** で切断したゲノム DNA は **HTLV-1** プロウィルスそれ自体が **PCR** の標的となるのを防ぐために、**Ligation** 反応後、**Mlu I** で切断したものを鋳型とした。

Inverse PCR の反応条件は以下の通りである。反応液は **50 μ l** で、各組成はプライマー (最終濃度 **0.2 μ M**)、**MgCl₂(2.5 mM)**、**dNTPS(0.4 mM)**、**10 \times LA buffer 5 μ l**、**LA Taq 0.5 μ l**、**template DNA (200 μ g)**である。**PCR** は、**94 $^{\circ}$ C**で 2 分インキュベーションすることで **DNA** を変成した後に **94 $^{\circ}$ C**で 30 秒、**68 $^{\circ}$ C**で 8 分というプログラムを **35** サイクル行った。

PCR 産物はエチジウムブロマイドを含んだ **1%アガロースゲル電気泳動**で解析した。全ての **PCR** は独立に **3** 回ずつ行った。

Pst I で切断したゲノムを鋳型とした場合、**primer 1** と **primer 2** を用いて **PCR** を行った。

Primer 1 5'- CAGCCCATTCTATAGCACTCTCCAGGAGAG -3'

Primer 2 5'- CAGTCTCCAACACGTAGACTGGGTATCCG - 3'

Xho I で切断したゲノムを鋳型とした場合、**primer 1** と **3** を用いて **PCR** を行った。

Primer 3 5'- AAGGAGTAGCGCGACAAGGGTGATTCCAGT -3'

EcoRI で切断したゲノムを鋳型とした場合、**primer 4** と **5** を用いて **PCR** を行っ

た。

Primer 4 5'-TGCCTGACCCTGCTTGCTCAACTCTACGTCTTTG-3'

Primer 5 5'-AGTCTGGCCCCTGACCTTTTCAGACTTCTGTTTC-3'

クローニングとサンガー法による HTLV-1 組み込み部位の決定

PCR 産物の精製に Quantum Prep™Freeze 'N Squeeze DNA Gel Extraction spin Columns を用い、pGEM-T easy vector(Promega)に ligation し、DH5α 大腸菌にトランスフォーメーション後、アンピシリン含有プレート上に展開し、12 時間以上 37°C でインキュベーションした。ランダムにピックアップしたコロニーを 2ml の 2×YT 培養液中に 37°C にて一晩培養し。プラスミド DNA をアルカリプレップ法で単離した。得られた DNA の 1/50 量を用い、EcoR I または Not I で消化後、アガロースゲル電気泳動にて目的の大きさのインサートの有無を確認した。同一のプレートから得られた同一サイズの複数のクローンでインサートの確認できたものについて塩基配列を決定した。Termo Sequence Cy5 dye terminator cycle sequencing kit (Amaersham Bioscience)を用いて PCR し、DNA をラベルした。その後、Gene Rapid (Amaersham Bioscience)を用いて塩基配列を解読した。

リアルタイム定量 PCR

プロウィルスのコピー数は Taqman 法によるリアルタイム PCR: the ABI PRISM 7000 sequence detection system (Applied Biosystems)を用いて計測した。定量 PCR には、PCR2 種類のプライマーセット（1つは HTLV-1 プロウィルス、もう1つは RNase P enzyme をコードする遺伝子）を使う multiplex PCR を行った。RNase P をコードする遺伝子に対するプライマーとプローブは Applide Biosystems から購入した。HTLV-1 プロウィルスの pX 領域に対するプライマーとプローブは以下の通りである。

Forward primer 5'-CGGATACCCAGTCTACGTGTT-3'

Reverse primer 5'-CAGTAGGGCGTGACGATGTA-3'

FAM-labeled probe 5'-CTGTGTACAAGGCGACTGGTGCC-3'

鋳型 DNA の対照検体として細胞あたり 1 コピーの HTLV-1 プロウィルスが組み込まれている TL-Om1 から抽出した染色体 DNA を使用した。

サザンブロットハイブリダゼーションデータベース

コホート研究により集積した PBMC 検体のうちプロウィルス量の多いものの一部についてはサザンブロットにより ATL 細胞のクローナリティを解析し、その結果をデータベース化されている。1 患者につき 2 回以上サザンブロット解析を

行っている患者について、Inverse PCR の結果と比較検討した。

次世代シーケンサーによる HTLV-1 組み込み部位の網羅的シーケンス
コホート研究より集積した PBMC 検体の HTLV-1 組み込み部位をイルミナ社、
次世代シーケンサー GA II、Hiseq を用いて index シーケンス及び 100bp ペアエ
ンドシーケンスを行った。

サンプルゲノムの断片化

新規クローナリティ解析システムで用いる検体はすべて sonication 装置 Covaris
(Covaris) を用いて断片化した。サンプル断片化条件は以下の通りである。

Micro Tube(6×16mm)Round bottom glass tube,AFA fiber with
Snap-Cap ,Covaris #520045

sample:10µg DNA in 100µl MQ

Covaris settings

Duty cycle.....	20%	5%
Intensity.....	5	3
Cycles per Burst.....	200	200
Time.....	5sec	90sec
Temperature.....	6-8°C	6-8°C

PCR 鋳型の作成

ランダムに断片化された DNA fragment に対して平滑化処理を施し、DNA 断片
末端に A 塩基を付加した後、アダプターをライゲーションした。実験条件は以
下の通りである。

sample(the product of Sonication).....	100µl	
MQ	43 µl	
10x DNA polymerase buffer	20µl	
ATP(10mM).....	20l(Takara#4041)(100mM	
)		
dNTP mix(25mM).....	8µl	(Invitrogen
#10297-018)		
T4 DNA polymerase(5U/µl).....	3µl	(Takara#2040A)
Klenow enzyme(5U/µl).....	1µl	(Takara #2140A)
T4 polynucleotide kinase(10U/□□l).....	5µl	(NEB #M0201L)
Total.....	200µl	

Incubate 20°C 30min
Clean up: PCR purification kit (Qiagen #28104)
Elute in 67µl MQ

A-tailing (1X rxn)

Sample(product of End repair).....67µl
NEB buffer2.....10µl
dATP(1mM).....20µl (Takara#4026)
klenow fragment exo- (15U).....3µl (NEB #M0212S)
Total.....100µl

Incubate 37°C 30min
Clean up: Qiaquick PCR purification kit (Qiagen #28104)
Elute in 60µl MQ

Adaptor ligation

Sample(product of A-taling).....30µl (split the product of A-taling into 2
tube of 30µl)
T4 DNA ligase buffer(10x).....5µl
Adaptor mix(25µM).....4µl
T4 DNA ligase enzyme(3U/µl).....5µl (NEB #M1801)
MQ.....6µl
Total.....50µl

Do the rxn in 500ml thin layer PCR tubes
Incubate 20°C 1~2hours
Mix total rxn 50+50µl in 1.5ml tubes
Clean up: MinElute PCR purification kit (Qiagen #28004)
Elute in 20µl MQ

過剰アダプターの除去

以下の条件でポリアクリルアミドゲルを用いて過剰アダプターを除去した。

30%acrylamide.....4.98ml
MQ.....8.415ml
10X TBE.....1.5ml
10% APS.....105µl

Total.....15ml
Then add 4.98µl TEMED
60min RT to solidify
Load 20µl sample in each well
Load 100bp DNA ladder
Running buffer 1x warm TBE
Pre run 100volt 15min
Run 160 volt 60 min
EtBr staining (100ml 0.5x TBE +5µl EtBr)10 min
Cut the gel from 200bp to 1000bp

アクリルアミドゲルから PCR 鋳型を切り出し以下の方法でゲル抽出を行った。

- Centrifuge 12000rpm 8min
- add (700µl) 1X LoTE buffer
- centrifuge 12000rpm 2 min
- Incubate at 65°C (water bath) 15 min
- Apply on the Colum(Costar Spin-X centrifuge TUBE Filters 0.22µm pore CA membrane ,#8160)
- Centrifuge 12000rpm 5min
- add 100µl LoTE
- centrifuge 12000rpm 5min
- keep the supernatant & throw out the Colum

ゲル抽出した PCR 鋳型は以下の条件でエタノール沈殿を行い、60µl 中 3µl をシーケンサーに持ち込み配列決定した。

- add(3M)NaAC (PH 5.2)(1:10 V:V) +1µl glycogen
- Add cold EtOH 100% (keep -20°C 30min)
- centrifuge 4°C 5 min
- add cold EtOH 70% 1ml
- centrifuge 4°C 10 min
- Air dry
- elute in 60µl MQ
- split in 3 tubes each 20µl for PCR

PCR

External PCR

DNA(product of size selection).....20µl
 Primer **F1**(10µM).....2µl
 Primer**R1**(10µM).....2µl
 10x bufferl5µl
 Accuprime taq high fidelity(5U/µl).....0.2µl(invitrogen #12346-094)
 MQ.....20.8µl
 Total.....50µl

PCR conditions

Initial denaturation94°C5min
 Denaturation94°C.....50sec } 25 cycles
 Combined annealing & extension ...68°C.....3min
 Final extension 68°C.....10min
 Hold@4°C

Ramping 1.9°C/sec

Mix the PCR product of the same samples 50+50+50µl total:15µl

Clean up:Qiaquick PCR purification kit (Qiagen #28104)

Elute in 300µl MQ

Nested PCR

DNA (2µl of The product of External PCR+18µl MQ).....20µl
 Primer **NFCB**(10µM).....2µl
 Primer**NP1F2**(10µM).....2µl
 10x bufferl5µl
 Accuprime taq high fidelity(5U/µl).....0.2µl(invitrogen #12346-094)
 MQ.....20.8µl
 Total.....50µl

PCR conditions

Initial denaturation 94°C5min
 Denaturation..... 94°C.....50sec } 30 cycles
 Combined annealing & extension.....68°C.....3min
 Final extension..... 68°C.....10min
 Hold@4°C

Ramping 1.9°C/sec

Clean up: Qiaquick PCR purification kit (Qiagen #28104)

Elute in 50µl MQ

Check the size & concentration of PCR product on **Bioanalyser**(Agilent DNA7500)&send for illumine sequencing

Required oligomers

Adaptors

Illu Long strand adaptor (86nt) HPLC

AACCGTTGCTAGGAGAGACCCAAGCAGAAGACGGCATACGAGATNNNNNNNGTGACT
GGAGTTCAGACGTGTGCTCTTCCGATCT

Illu Short strand adaptor (27nt) HPLC

P-GATCGGAAGAGCGTTTTTTTTTCAAAA

Primers

F1 (20nt)EX oligo

TACCGGCGACTCCGTTGGCT

R1 (28nt)EXoligo

CGAAGAGTAACCGTTGCTAGGAGAGACC

NFCB(21nt) EXoligo

CAAGCAGAAGACGGCATACGA

NP1F2 (81nt) platinumoligo

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGA
TCTCCAGCGACAGCCCATCCTATAGC

HTLV-1 組み込み部位の同定

HTLV-1 の組み込み部位を含む read を探索するために、BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>) プログラムを利用した。HTLV-1 LTR 領域と LTR と隣接する gag 領域をリファレンス配列として用いた。BLAST を用いて read を mapping し、LTR 配列の末端配列を read の 5'末端に含む read を

HTLV-1 組み込み部位配列とした。LTR 配列は HTLV-1 ゲノムに 5'側、3'側に 2 箇所存在する。HTLV-1 組み込み部位を増幅する際、HTLV-1 5'側 LTR も増幅される (Fig 10)。HTLV-1 5'側 LTR が増幅された場合、ヒトゲノム配列は PCR 産物に含まれないため、クローン検出には利用できない。HTLV-1 5'側 LTR 由来の read を除去するため、LTR 配列と隣接する gag 配列を合わせたリファレンスゲノムを用いた。read 配列に gag 配列が含まれている場合、その read は HTLV-1 5'側 LTR 由来と判断し、その後の解析には利用しない。LTR 末端配列のみ含まれる read (read1)、及び pair read(read2)をヒトゲノムへ mapping した。

ヒトゲノム中の HTLV-1 組み込み部位を探索するために Genomatix 社が提供する Genomatix Mining Station を利用し read をヒトゲノム(hg 19)へ mapping した。HTLV-1 組み込み部位を含む read1 は HTLV-1 LTR 配列を除去してヒトゲノムへ mapping を行った。また、read1、read2 それぞれ 5 カ所の mutation を許容して mapping を行った。unique に mapping された read の内、read1 と read2 の間の距離が 1kbp 以内の read1 をもとにクローンのゲノム組み込み部位を同定した。

DNA shear site の同定

DNA shear site を含む read は read1 のペア read の read2 である。read2 は HTLV-1 LTR 領域は含む場合と含まない場合が考えられる。ゲノムは約 1kbp に断片化されている。ゲノムはランダムに断片化されており、HTLV-1 の 3' LTR 領域が断片配列のどこに存在するかは予測できない。3'側 LTR の近傍で断片化され、1kbp 断片に HTLV-1 配列を LTR 配列以外ほとんど含まない場合は、PCR 産物は 1kbp 前後になるため、shear site を含む read2 に LTR 配列は含まれない。一方、1kb 断片の大部分に HTLV-1 の配列を含む場合、PCR 産物は非常に短くなり、shear site を含む read2 に HTLV-1 LTR 配列が含まれてくる(Fig 11)。read2 に LTR 配列を含んだ場合、LTR 配列を取り除きヒトゲノムに mapping する必要がある。100bp 全長をヒトゲノムに mapping 後、unique に mapping されなかった read は BLAST を用いて HTLV-1 LTR 配列に mapping した。read2 の末端に LTR 配列を含む read を抜き出し、LTR 配列を取り除いた read2 を再度ヒトゲノムへ Genomatix Mining Station を用いて mapping した。BLAST はマッチする LTR 配列が 10bp 以下の場合、LTR を含む配列を検出することはできない。この点を考慮し、BLAST で検出できなかった read2 は全て、3'末端を 10bp 除去後、ヒトゲノムへ mapping した。それぞれ unique に mapping された read2 から shear site を同定した。

HTLV-1 感染細胞クローンサイズの計測

プログラム言語 **perl** を用いてクローンサイズを測定するプログラムを作成した。クローンサイズを測定するためには各クローンを構成する **sister cell** の数を測定する必要がある。**Sister cell** 数は **DNA** を断片化する際得られる **shear site** 及び、アダプターにデザインした **8bp** のタグ配列をもとに測定する。**Shear site** の情報は **mapping** された **read2** のゲノム位置情報から得られる。**Read1** のゲノム位置が同一で、**read2** のゲノム位置が異なる場合異なる **sister cell** 由来 **read** とみなす。また、**read2** のゲノム位置が同一でもタグ配列が異なる場合は異なる **sister cell** 由来 **read** とみなす。**Read2** が同一ゲノム位置に **mapping** され、同一タグ配列を持つ場合、**PCR duplicate** とみなし **sister cell** を測定した。(Fig 8)

HTLV-1 組み込み部位の特徴付け

次世代シーケンサーから得られた **HTLV-1** の組み込み部位のゲノムアノテーション解析を **Genomatix** 社が提供する **Genomatix Genome Analyzer** を利用した。

実験結果

Inverse PCR による ATL 細胞検出系の確立

ATL 細胞はこれまで血清中の LDH 値、及び可溶性 IL-2 レセプターの値が ATL 細胞数と相関するため、これらの値を基準として ATL 細胞数を推定してきた。しかし、これらの方法は直接的に ATL 細胞を検出し、定量しているわけではない。また、末梢血における ATL 細胞は「花細胞=Flower cell」と呼ばれる形態をとることが知られている。しかし、全ての ATL 細胞がこの形態を示すわけではない。以上のような事情により ATL 細胞を直接的に定量する方法がないのが現状である。そこで我々は Inverse PCR を用いた ATL 細胞検出系を構築し、JSPFAD コホートと研究により集積された患者検体 20 名分（計 60 検体）の ATL 細胞の検出を試みた。具体的には末梢血ゲノム DNA を鋳型とし、Inverse PCR による解析を行った（患者の病型は Table 1 を参照）。Inverse PCR の結果を確認するために Pst I、EcoR I の 2 種類の制限酵素を用いる解析を行った。また Inverse PCR を 1 検体につき 3 回行うことで PCR tube 間での反応の差を検証した。Pst I、EcoR I どちらの場合も Inverse PCR での主な PCR 産物=メインバンドを ATL 細胞クローン由来 PCR 産物とみなした。以下に結果をしめす

1. Pst I、EcoR I 両方でメインバンドが確認できた検体
検体 No. 1、3、7、12、13、14、18
2. Pst I のみでメインバンドが確認できた検体
検体 No. 8、11、16、17
3. EcoR I のみでメインバンドが確認できた検体
検体 No. 5、6、10、15、19
4. Pst I、EcoR I 両方でメインバンドが確認できなかった検体
検体 No. 20

20 名中 19 名の患者検体からメインバンドを検出できた。メインバンドが検出できた患者検体のうち 3 名についてはメインバンドの位置が時間経過に伴い異なっていた。また、メインバンドの濃淡はリアルタイム PCR を用いた感染細胞数（ウィルスロード）と一致しなかった。また、ATL クローン以外のマイナークローンは数種類しか検出できなかった。上記の結果より Inverse PCR 法は ATL 細胞を直接検出できるが、クローンを網羅的に検出できずクローナリティ解析に用いることはできないと示唆された。

HTLV-1 組み込み部位を用いたクローンの検出

HTLV-1 はレトロウイルスで、1 細胞に 1 か所 HTLV-1 組み込み部位を持つ。1 感染細胞が増殖しクローンを構成する。我々は HTLV-1 のヒトゲノム組み込み部位を調べることでクローン検出を試みた。HTLV-1 感染キャリアにおける HTLV-1 感染率は 2~3% (100PBMC 中) と低いため、HTLV-1 組み込み部位を増幅する必要がある。そこで我々は **Nested splinkerette PCR** 法を用いて HTLV-1 組み込み部位を増幅し、次世代シーケンサーを用いて PCR 産物の配列決定を行った (32) (Fig12)。既存の HTLV-1 クローン検出法は実験条件を新規検出系と統一するため、制限酵素を用いた **Nested splinkerette PCR** 法で増幅した PCR 産物をゲル電気泳動で確認した (Fig13 A)。バンドとして区別できる PCR 産物をクローンとみなした。クローン検出は P1(80%)、P2(10%)、P3(1%) の 3 検体を利用した。()はウイルスロードを表す。

P1 は ATL 患者の検体である。P1 検体の場合、**major clone** がゲル電気泳動、次世代シーケンサーを用いた新規 HTLV-1 クローン検出系両方で 2 つ確認できた。しかし、既存検出系は **major clone** 以外は検出できなかった。新規検出系は ATL クローン候補以外にも 50 クローン検出できた。他 2 検体も同様に新規検出系は既存検出系と比較して非常に多くのクローンを検出することができた。上述の結果より、新規検出系は HTLV-1 感染細胞クローンを網羅的、特異的に検出することができたと考えられる。

Genome Shear site を利用したクローンサイズの計測

ATL クローン候補を早期に発見するためにはクローンサイズを正確に計測する必要がある。クローンサイズを計測するためにはクローンを構成する **sister cell** 数を計測する必要がある。**Sister cell** とは 1 種類の感染細胞から増殖した異なる感染細胞のことを表す。各クローンは HTLV-1 のヒトゲノムへの組み込み部位で判別したが、**Sister cell** は同一組み込み部位を持つ配列から判別しなければならない。2011 年 1 月に英国のグループより、ゲノム断片化処理で生じる **DNA shear site** を利用した **sister cell** の判別法が発表された(32)。そこで我々も同様の方法を用いて **Sister cell** の判別を行い、クローンサイズを計測した (Fig 13-B)。

P1 検体は既存クローン検出システムを用いた場合、**Major clone** が 2 つ確認できる。新規クローン検出システムも同様に **Major clone** が 2 つ確認できることからクローンサイズは確からしいと判断できる。しかし、ATL 患者検体は **Major clone** が 1 つのモノクローナルパターンがほとんどである。つまり、既存システム、及び、新規クローン検出システムどちらもクローンサイズを正しく計測できていない可能性がある。P 2、3 はキャリアの検体である。キャリアの多くは **Major clone** の割合が小さいことが既存研究で知られている。既知の情報と一致

するため、正確にクローンサイズを計測できた可能性が高い。

タグ配列を用いた新規クローナリティシステムの開発

英国のグループが用いているクローンサイズの計測方法は正確性に問題点がある。彼らは DNA 断片化処理の際生じる断片化サイト (DNA shear site) を用いてクローンを構成する sister cell を計測している。ATL のように 1 種類のクローンがモノクローナルに増殖した場合、DNA shear site では sister cell を判別しきれない可能性がある。理由を以下に述べる。DNA shear site はシーケンス効率を最も高くするために、ゲノムを約 300bp~700bp に断片化する。その結果、shear site は最大約 400 に留まり、sister cell は 400 個以上判別できない。つまり ATL のモノクローナルなクローンの場合、400 以上の Sister cell で構成されている可能性が高いにも関わらずクローンを構成する感染細胞数は 400 に誤認識される恐れが生じる (Fig 14-A)。またクローンを構成する Sister cell が 50 以上になると、偶然同一 shear site を生じる確率が 95%以上となる (Fig14-B)。つまり shear site のみでは正確なクローンサイズの計測は不可能である。この問題点を改善するために我々は 8bp のランダムなタグ配列をライゲーションするアダプター配列に追加した独自の検出系を開発した (Fig15)。8bp のランダムな配列を持つタグは 4 の 8 乗種類、つまり 65536 通りバリエーションがある。DNA shear site にこのランダムなタグ配列を持つアダプターをライゲーションすると、HTLV-1 感染細胞ごとに異なるタグ配列を含む PCR 鋳型ができる。この鋳型をもとに PCR を行い、配列を決定する。タグ配列が異なれば DNA Shear site が同じでも異なる Sister cell とみなすことで $400 \times 65536 = 26214400$ 通り細胞を区別できる (Fig15)。

過剰アダプターの除去

Nested splinkerette PCR 法を用いる際、残存アダプター配列がプライマーとして機能する可能性が高い (Fig 16)。残存アダプターがプライマーとして機能すると、タグ配列と Sister cell の関連が全くなり、クローナリティ解析ができない。この問題点を解決するため、ポリアクリルアミドゲルに PCR 鋳型を泳動し、アダプター配列を除去した後、PCR を行った (Fig 17)。ポリアクリルアミドゲルによるアダプター配列除去によりタグ配列の濃縮が確認されクローナリティ解析が可能となった。

タグ配列の有効性の検証

急性型 ATL 検体 (P4) を用いてタグ配列の有効性について検証した (Fig18)。下側の棒グラフが Shear site を利用した場合のクローナリティ、上側の棒グラフ

がタグ配列、**shear site** 両方を利用したクローナリティ解析結果を示した。**shear site** のみ利用した場合、**shear site**、タグ配列を利用した場合どちらも **Major clone** が検出された。**Major clone** はどちらも同じゲノム組み込み部位であったため同一クローンと考えられる。**Major clone** のクローナリティはタグ配列の有無により全く異なる結果が得られた。**Shear site** から得られる **Major clone** の **Sister cell** 数は全 HTLV-1 感染リンパ球 3256 中 912 (28.0%) であった。全 HTLV-1 感染細胞数は 10,000 リンパ球中に存在する HTLV-1 感染リンパ球をウィルスロードより算出した。タグ配列を利用することで同じ **shear site** を持つシーケンスリードを **sister cell** の違いにより区別した結果、**Major Clone** の **Sister cell** を全 HTLV-1 感染リンパ球 3256 中 2621(80.5%) 得ることができた。**Second major clone** も **Shear site** を利用した場合、**Shear site**、タグ配列両方を利用した場合どちらも同一の **clone** であった。**Shear site** から得られる **Second major clone** の **Sister cell** は 102(3.1%) であった。タグ配列を利用し、**sister cell** を区別した結果 **Second Major clone** の **Sister cell** は 32(1.0%) であった。上記の結果より、タグ配列を用いることで **Major clone** のクローナリティが大きく変化することが明らかとなった。急性型 ATL はモノクローナルな HTLV-1 の組み込みが確認されることがわかっている。タグ配列を利用したクローナリティがより正確である可能性が高いと考えられる。

急性型 ATL 検体以外にくすぶり型 ATL2 検体 (P5(VL:7%), P6(30%))、無症候性キャリア 1 検体(VL:7%) についてタグ配列の有無によるクローナリティの違いについて検証した (Table 1)。無症候性キャリアの検体(P5)のクローナリティはタグがある場合は **Major clone**、**Second major clone** が 10%、5% であったのに対し、タグない場合は 5%、4% でタグの有無により大きな違いはなかった。くすぶり型検体 (P6) は **Major clone**、**Second Major clone** の割合は **Shear site** のみを利用した場合 29%、18% であったのに対し、タグ配列、**shear site** 両方を利用した場合 22%、21% とほとんど違いがなかった。しかし、**third major clone** の割合はタグの有無により、16% (**Shear site** のみ)、5%(タグ配列、**Shear site**) と違いが生じた。くすぶり型検体 (P7) の場合、**Major clone** の割合が 26%(**shear site** のみ)、48%(タグ配列、**Shear site**) と違いが生じた。

クローナリティの再現性の検討

新規クローナリティ解析システムから得られたクローン、クローナリティの再現性について検討した。P4 (急性型 ATL VL:7%)、P6(くすぶり型 ATL VL:30%) を検証サンプルとして用いた。2 検体を個別に 2 回シーケンスを行い、得られたクローン、クローンサイズを比較した (Fig19)。

P4 サンプルは 398 クローン中、198 クローンが一致した。クローンサイズは上

位 10 クローン中 8 クローンが一致した。また上位 2 クローンは順位も一致していた(Fig 19-A)。

P6 サンプルは 52 クローン中、25 クローンが一致した。クローンサイズは上位 10 クローン中 7 クローンが一致した。上位 2 クローンは順位を一致していた(Fig 19-B)。

2 検体いずれも **Major clone** の検出に成功しており再現性があると考えられる。

病型の異なる ATL のクローナリティ解析

ATL の病型は急性型、慢性型は早急な治療を必要とする **Aggressive** タイプと経過観察を行う慢性型、くすぶり型の **Indolent** タイプに分類することができる。くすぶり型は異常リンパ球が 5%以上存在する以外は無症候性キャリアと臨床状態がない場合が多い。しかし、無症候性キャリアを比べ明らかに予後が悪いことが知られている。そこで異なる病型の ATL 及び、同レベルのウィルスロードをもつ無症候性キャリアの検体とクローナリティの違いについて検証した。まず、ウィルスロードが同レベルの異なる病型を示す ATL 患者のクローナリティを比較した(Fig 20 A,B)。ウィルスロードはそれぞれ 31.15%、32.56%である。ウィルスロードより算出した 10,000 リンパ球中に存在する全 HTLV-1 感染細胞数は 3115、3256 である。クローンはそれぞれ、くすぶり型検体は 220 クローン、急性型検体は 325 クローン得られた。くすぶり型、急性型 ATL 検体から得られた **Major clone** のクローナリティはそれぞれ 1503 (48.2%)、2621(80.5%)であった。**Second Major clone** のクローナリティはそれぞれ 812 (26.0%)、102 (1%) であった。急性型 ATL の場合、**major clone** が全 HTLV-1 感染細胞の大半を占めているのに対してくすぶり型の場合その割合が少ない。また、急性型 ATL は **second major clone** の割合は小さいのに対して、くすぶり型 ATL の場合その割合が大きい。上述の結果より病型によりクローナリティが異なる可能性が示唆される。

次にウィルスロードが同程度のくすぶり型患者、無症候性キャリアの HTLV-1 感染細胞クローナリティを比較した(Fig 20 A,C)。ウィルスロードはそれぞれ 9.01%、7.57%であった。ウィルスロードより算出した 10,000 リンパ球中に存在する全 HTLV-1 感染細胞数は 901、757 である。くすぶり型検体から 52 クローン、無症候性キャリア検体から 827 クローン得られた。**Major clone** のクローナリティはそれぞれ 263(27%)、83(10%)であった。また **Second Major clone** のクローナリティは 161 (16%)、38 (5%) であった。2 検体を比較すると、**Major clone** のクローナリティが異なることが確認できた。ウィルスロードが高いくすぶり型 ATL のクローナリティとウィルスロードの低いくすぶり型 ATL のクローナリティは 2~3 クローンの全感染細胞中に占める割合が高いという点で類似し

ていることがわかった。

Major clone のヒトゲノム組み込み部位の特徴

ATL4 検体の Major clone の組み込み部位を検証した (Fig 21)。

P1 の Major clone の組み込み部位は GDPD2 と DLG3 遺伝子間であった (Fig21-A)。P1 の Second Major clone の組み込み部位は BMPER 遺伝子と transcript ENST00000446484 の間であった (Fig21-B)。P4 の Major clone の組み込み部位は HDX 遺伝子のイントロンであった (Fig21-C)。P5 の Major clone の組み込み部位は PIGG 遺伝子と FLJ41130 遺伝子間であった (Fig21-D)。また、P5 の Second Major clone の組み込み部位は周辺 30kbp に遺伝子の存在しない intergenic なゲノム位置であった (Fig21-E)。

P6 の Major clone の組み込み部位は RNF111 のイントロンであった (Fig 20-F)。また、P6 の Second Major clone の組み込み部位は周辺 30kbp に遺伝子の存在しない intergenic なゲノム位置であった (Fig21-G)。

4 検体の Major clone、second major clone のゲノム組み込み部位はイントロンが 2 クローン (7 クローン中)、intergenic な組み込み部位は 5 クローン (7 クローン中) と intergenic な領域に集中していた。

HTLV-1 ヒトゲノム組み込み部位の特徴

7 検体から得られた全 HTLV-1 ヒトゲノム組み込み部位の特徴を検証した。まず、7 検体から得られた全組み込み部位のヒトゲノムアノテーション (エクソン、イントロン、intergenic、プロモータ) を確認し、各ゲノムアノテーションの割合を算出し、ヒトゲノムアノテーションの割合を比較した (Fig 22)。比較スコアは enrichment score として表した (Table 2)。プロモータ領域は各 transcript の第一エクソン上流 500bp、下流 100bp とした。全 5948HTLV-1 組み込み部位のゲノムアノテーションを解析した結果、プロモータ領域に組み込み部位が集中し、エクソン領域に少ない可能性が高いことが確認できた。ATL と無症候性キャリア検体を分類し、HTLV-1 組み込み部位を解析した結果、ATL4 検体の HTLV-1 組み込み部位はエクソン領域への組み込み頻度が他ゲノムアノテーションより低い結果が得られた (Fig.22-A)。また、プロモータ領域に組み込み部位が集約されていることが確認できた。無症候性キャリアの 3 検体は promoter 領域への集約が確認できた (Fig.22-B)。HTLV-1 組み込み部位を持つプロモータが制御する 202 種類の下流遺伝子について MeSH term を用いた関連解析を行った (Table 3)。その結果、遺伝子群がリンパ球関連 term と関連性が高いことが確認できた。また、120 遺伝子が Unigene データベースから確認できた。プロモータ領域とクローン構成感染細胞数との相関は見られなかった。

ヒトゲノム組み込み部位の方向性検証

HTLV-1 はヒトゲノムの 5'側、3'側の方向性を考慮せず組み込まれる。つまり、ヒトゲノムと逆向きに組み込まれるウイルスも存在する (Fig. 24)。そこでウイルスゲノムが組み込まれる方向性について検証した (Table 4)。

まず、ATL4 検体についてヒトゲノム組み込み部位の方向性を検証した。全クローンの組み込み部位の方向性を検証した結果、組み込み部位の方向性に偏りは確認できなかった。次にクローンサイズを 3 つに分類し、組み込み部位の方向性について検証したところ、Major clone、Middle major clone は僅かに plus strand に組み込まれる頻度が高いことが確認できた。

次に Carrier の 3 検体についてヒトゲノム組み込み部位の方向性を検証した。キャリアは Major clone を持たないため、クローンサイズの分類指標を変更した。キャリアの検体も ATL 検体同様、Major clone、Middle major clone は僅かに plus strand に組み込まれる頻度が高いことが確認できた。ATL、およびキャリアどちらもほとんどのクローンが minor clone であることが確認できた。

考察

まず **Inverse PCR** 法を用いた **ATL** クローン検出系を開発した。制限酵素を複数利用することで **20** 検体中、**19** 検体で **ATL** クローンの候補を検出することができた。また、ゲル電気泳動結果から得られる **PCR** 産物のバンドの濃淡から **ATL** クローンの増減を相対的に判断することは可能であった。しかし、**Inverse PCR** 法を用いる場合、ヒトゲノムの制限酵素サイトが **HTLV-1** 組み込み部位の近傍にあるクローンしか検出できない、また、感染細胞クローン全体を捉えることができないため絶対的なクローンサイズを測定するクローナリティ解析を行うことはできないという問題点が残る。クローナリティ解析は早期に **ATL** クローンの候補を検出する、また、**ATL** の病型を分類する指標として利用が期待できる。そこで我々は次世代シーケンサーを用いてレトロウィルスのヒトゲノム組み込み部位の配列を決定し、クローンサイズを測定できるシステムを開発した。クローナリティ解析システムで最も重要なポイントは正確にクローンサイズを計測することである。**2011** 年 **1** 月に英国のグループより発表されたクローナリティ解析システムは各細胞がもつ **DNA shear site** を用いてクローンサイズを測定していた。しかし、**DNA shear site** は最大 **400** サイト程度に止まる。**ATL** クローンなど特定のクローンがモノクローナルに増殖した場合、別細胞にもかかわらず、同一の **Shear site** を生じる頻度が上昇するため、正確なクローンサイズの計測ができない。英国のグループのこの問題点を認識しており、統計的に補正した結果を示している。

我々は統計的な補正を用いず、直接感染細胞を個別に識別できるシステムの開発を目指した。具体的にはランダムな塩基配列 **8bp** のタグ配列を追加したアダプターを用いて感染細胞の識別を目指した。タグのバリエーションは **4** の八乗 (**65,536**) 種類ある。タグ配列と **DNA shear site** を組み合わせると **2600** 万以上の細胞を識別できる。本システムに用いる白血球数は最大約 **100** 万細胞であるため、全感染細胞を網羅することが可能であると考えられる。

タグ配列の有無がクローンサイズを大きく変化させることが確認できた。特に急性型 **ATL** 検体 (**P4**) の **Major clone** のクローンサイズに大きな違いが生じた。急性型 **ATL** の多くはモノクローナルなクローンがサザンブロットにより検出されることが知られている。タグ配列はモノクローナルなクローンサイズの測定に最も影響することが示唆される。**P1** 検体は **80%** と高いウィルスロードにも関わらず、**2** つの **Major clone** が確認された。**P1** 検体はタグ配列情報を持たないため、**shear site** を利用したクローナリティ解析を行っている。タグ配列を用いて再度解析すると異なる結果が得られるかもしれない。

新規クローナリティ解析システムから得られたクローンサイズの再現性について

て ATL2 検体を用いて検証した。同一サンプルに対して個別に PCR を行い、クローナリティを解析した結果、2 検体とも上位 10 クローンのクローンサイズ再現性が確認された。再現性を確認した 2 検体は同一検体であるが、ゲル電気泳動を用いた過剰アダプター除去処理の有無が異なる。過剰アダプターを除去しない場合、タグ配列を用いたクローナリティ解析を行うことはできない。過剰アダプターを除去したサンプル (P4-1,P6-1) はタグ配列、Shear site を用いたクローナリティ解析を行い、過剰アダプターを除去していないサンプル (P4-2, P6-2) は shear site のみを利用したクローナリティ解析を行い、各クローンサイズを比較した。クローンサイズの測定方法が異なるため、直接サンプル間のクローンサイズを比較することはできないため、サンプル内のクローンサイズの大きい 10 クローンを抽出し比較した。その結果、P4 サンプルは 10 クローン中 7 クローン、P6 サンプルは 10 クローン中 7 クローンが同一クローンであったため、我々が開発したクローナリティ解析システムのクローンサイズの再現性は高いと考えられる。

各検体の HTLV-1 組み込み部位を抽出し、ゲノムアノテーションと比較し組み込み部位の特徴について検証した。これまで HTLV-1 の組み込み部位はランダムにヒトゲノムに組み込まれるという意見がほとんどであるが、AT rich な配列に組み込まれる頻度が高い(38)、ATL クローンは転写開始点近傍に組み込まれる頻度が高い(39)など様々な論文が報告されており結論は出ていない。しかし、どの論文も数十の組み込み部位を検証したのみであり、網羅的な HTLV-1 組み込み部位の検証はこれまで不可能であった。そこで本研究で用いた 7 検体のゲノム組み込み部位の特徴について検証した。7 検体のゲノム組み込み部位はプロモータ領域の頻度が高いことが確認された。一方、エクソンへの組み込み頻度が低いことも確認された。これらの結果はクロマチン構造が転写制御により不安定であり、HTLV-1 が容易に組み込み易い可能性を示唆している。HTLV-1 以外のレトロウィルスの宿主ゲノムへの組み込み部位網羅的解析は既にいくつか報告されている(35)。MLV は転写開始点近傍に組み込み部位が集約し、HIV はランダムな組み込みが報告されている。HTLV-1 と HIV は同じレンチウイルスに属するため、HIV の組み込みの特徴に類似している可能性が高い。しかし、HTLV-1 の組み込み部位の特徴について結論づけるにはより多くの検体を検証する必要がある。また、クローンサイズを分類し組み込み部位のゲノムアノテーションを検証する必要と考えられる。

ATL は急性型、リンパ腫型、慢性型、くすぶり型の 4 つに分類される。慢性型、くすぶり型は臨床的病状を伴わない場合が多いため経過観察をとるが、急性転化を生じるなど無症候性キャリアと比較すると明らかに予後が悪いことが知られている。くすぶり型は特に異常リンパ球が 5%以上という分類指標以外無症候

性キャリアと同じであるため、診断が難しいとされている。そこで今回確立した新規クローナリティ解析システムを用いてくすぶり型 2 検体、急性型 1 検体、無症候性キャリア 1 検体を用いてクローナリティを比較検討した。その結果 4 検体それぞれ Major clone、Second major clone のクローナリティが異なることが確認できた。急性型 ATL は Major clone が全体の 80%以上を占めていたのに対して、くすぶり型 2 検体は Major clone、Second major clone が大きな割合を占めていた。くすぶり型 2 検体も 2 つの Major clone があるという点では類似していたが、そのクローナリティは異なっていた。2 検体のウィルスロードは 30%と 7%で異なる。30%ウィルスロード検体 (P5) の方が Major clone の割合が大きく、急性型の Major clone の割合に近づいている。この結果から、ウィルスロードが高く、Major clone の割合が高い検体 (P5) は急性転化する可能性が高いと判断できる。くすぶり型検体を Inverse PCR を用いて検証した場合、クローンのモノクローナルな組み込みは明確に確認できていない。複数のクローンが確認できるポリクローナルな状況に近いことが確認されている。Inverse PCR の結果からもくすぶり型 ATL は複数の Major clone が存在している可能性が高いことを支持していた。

ATL 発症リスクファクターとしてウィルスロードが用いられている。ウィルスロードが 4%以上のキャリアからのみ ATL を発症することが JSPFAD コホート研究より明らかとなっている。しかし、4%のキャリアの多くは無症候性のままでいる場合が多い。今回我々が確立した新規レトロウィルスクローナリティ解析システムを用いることで低ウィルスロードの検体のクローナリティ解析が可能となった。その結果予後の悪い急性型 ATL は単一クローンのモノクローナルな増殖が確認できた。比較的症状の軽いくすぶり型検体はウィルスロードに関わらず複数のクローンの増殖が確認された。無症候性キャリアは特徴的に増幅したクローンは検出されなかった。これらの結果から HTLV-1 感染キャリアの ATL 発症リスクファクターはウィルスロードと各クローンのクローナリティであることが示唆される。つまり単一クローンのモノクローナルな増殖が HTLV-1 感染キャリアの予後不良因子である可能性が高い。Inverse PCR を用いた ATL 細胞検出システムを利用した際、複数の検体で ATL クローンが経時的に変化することが確認された。変化した際、病態も慢性型から急性転化、慢性型から自然寛解と変化していた。

クローナリティ解析より得られた ATL の病型分類、予後予測をモデル図として表した(Fig25)。ウィルスロードとクローナリティ解析を組み合わせることで早期 ATL 発症予測また ATL 予後予測、ATL 病型分類に貢献することが可能になると示唆される。

謝辞

本研究においてご指導賜りました東京大学新領域創成科学研究科メディカルゲノム専攻病態医療科学分野教授 渡邊 俊樹先生に心より感謝の意を表します。また、終始ご指導ご鞭撻を賜り副査として助言をいただきました東京大学医科学研究所 ヒトゲノム解析センター 機能解析イン・シリコ分野教授 中井謙太先生に深く感謝の意を表します。

本研究を遂行するにあたり、次世代シーケンサーを用いたデータを提供していただき、また副査として助言を戴いた東京大学新領域創成科学研究科メディカルゲノム専攻ゲノム制御医科学分野教授 菅野純夫先生、及び准教授 鈴木穰先生に心より感謝の意を表します。

並びに副査として助言を戴くとともに本論文の細部にわたりご指導いただいた東京大学医科学研究所 癌・細胞増殖部門人癌病因遺伝子分野教授 村上善則先生、東京大学医科学研究所附属病院 血液腫瘍内科准教授 内丸 薫先生に深く感謝の意を表します。

本研究を遂行するに当たり始終ご指導くださいました東京大学医科学研究所ヒトゲノム解析センター 機能解析イン・シリコ分野特任助教 山下 理宇先生、東京大学新領域創成科学研究科メディカルゲノム専攻病態医療科学分野准教授 佐藤 均先生、及び、助教 中野 和民先生に深く感謝の意を表します。

本研究を遂行するにあたり、サンプル作成に協力いただき、また実験プロトコルの完成に大きく寄与いただいた東京大学新領域創成科学研究科メディカルゲノム専攻修士二年 **Sanaz Firouzi** 氏に深く感謝いたします。

引用文献

1)

Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma.

Poiesz BJ, Ruscetti FW, Gazdar AF, Bunn PA, Minna JD, Gallo RC.

Proc Natl Acad Sci U S A. 1980 Dec;77(12):7415-9.

2)

A retrovirus associated with human adult T-cell leukemia: in vitro activation.

Hinuma Y, Gotoh Y, Sugamura K, Nagata K, Goto T, Nakai M, Kamada N, Matsumoto T, Kinoshita K.

Gann. 1982 Apr;73(2):341-4.

3)

Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease.

Yoshida M, Miyoshi I, Hinuma Y.

Proc Natl Acad Sci U S A. 1982 Mar;79(6):2031-5.

4)

Global epidemiology of HTLV-I infection and associated diseases.

Proietti FA, Carneiro-Proietti AB, Catalan-Soares BC, Murphy EL

Oncogene. 2005;24:6058–6068.

5)

Human T-lymphotropic virus type II and neurological disease.

Araujo A, Hall WW.

Ann Neurol. 2004;56:10–19.

6)

Monoclonal integration of human T-cell leukemia provirus in all primary tumors of adult T-cell leukemia suggests causative role of human T-cell leukemia virus in the disease.

Yoshida M, Seiki M, Yamaguchi K, Takatsuki K.

Proc Natl Acad Sci U S A. 1984 Apr;81(8):2534-7.

7)

Treatment of adult T-cell leukemia-lymphoma with a combination of interferon alfa and zidovudine.

Gill PS, Harrington W Jr, Kaplan MH, Ribeiro RC, Bennett JM, Liebman HA, Bernstein-Singer M, Espina BM, Cabral L, Allen S, et al.

N Engl J Med. 1995 Jun 29;332(26):1744-8.

8)

Diagnostic criteria and classification of clinical subtype of adult T-cell leukemia-lymphoma

Shimoyama M.

a report from the Lymphoma Study Group (1984-87). Br J Haematol 1991;79(3):428-437.

9)

A new G-CSF-supported combination chemotherapy, LSG15, for adult T-cell leukaemia-lymphoma

Yamada Y, Tomonaga M, Fukuda H, Hanada S, Utsunomiya A, Tara M, Sano M, Ikeda S, Takatsuki K, Kozuru M, Araki K, Kawano F, Niimi M, Tobinai K, Hotta T, Shimoyama M.

Japan Clinical Oncology Group Study 9303. Br J Haematol 2001;113(2):375-382

10)

Allogeneic hematopoietic stem cell transplantation provides sustained long-term survival for patients with adult T-cell leukemia/lymphoma.

Fukushima T, Miyazaki Y, Honda S, Kawano F, Moriuchi Y, Masuda M, Tanosaki R, Utsunomiya A, Uike N, Yoshida S, Okamura J, Tomonaga M.

Leukemia 2005;19(5):829-834.

11)

VCAP-AMP-VECP compared with biweekly CHOP for adult T-cell leukemia-lymphoma:

Kunihiro Tsukasaki, Atae Utsunomiya, Haruhiko Fukuda, Taro Shibata, Takuya Fukushima, Yoshifusa Takatsuka, Shuichi Ikeda, Masato Masuda, Haruhisa Nagoshi, Ryuzo Ueda, Kazuo Tamura, Masayuki Sano, Saburo Momita, Kazunari Yamaguchi, Fumio Kawano, Shuichi Hanada, Kensei Tobinai, Masanori Shimoyama, Tomomitsu Hotta and Masao Tomonaga

Japan clinical oncology group study JCOG 9801. J Clin Oncol 2007;25(34):5458-5464

12)

Chemotherapy of ATL. Adult T-cell Leukemia.

Shimoyama M; Takatsuki K

Oxford. Oxford University Press; United Kingdom; 1994. p. 221-237

12)

Real-time polymerase chain reaction for quantification of HTLV-1 proviral load: application for analyzing aberrant integration of the proviral DNA in adult T-cell leukemia.

Kamihira S, Dateki N, Sugahara K, Yamada Y, Tomonaga M, Maeda T, Tahara M.

Int J Hematol. 2000 Jul;72(1):79-84.

13)

Polyclonal integration of HTLV-I proviral DNA in lymphocytes from HTLV-I seropositive individuals: an intermediate state between the healthy carrier state and smouldering ATL.

Yamaguchi K, Kiyokawa T, Nakada K, Yul LS, Asou N, Ishii T, Sanada I, Seiki M, Yoshida M, Matutes E.

Br J Haematol. 1988;68(2):169-74. doi: 10.1111/

13)

High human T cell leukemia virus type-1(HTLV-1) provirus load in patients with HTLV-1 carriers complicated with HTLV-1-unrelated disorders.

Sasaki D, Doi Y, Hasegawa H, Yanagihara K, Tsukasaki K, Iwanaga M, Yamada Y, Watanabe T, Kamihira S.

Virology. 2010 Apr 28;7:81.

14)

Proviral status of HTLV-1 integrated into the host genomic DNA of adult T-cell leukemia cells.

Kamihira S, Sugahara K, Tsuruda K, Minami S, Uemura A, Akamatsu N, Nagai H, Murata K, Hasegawa H, Hirakata Y, Takasaki Y, Tsukasaki K, Yamada Y

Clin Lab Haematol. 2005;27:235-41. doi: 10.1111/j.1365-2257.2005.00698.x.

15)

An ATL cell line with an IgH pseudo-rearranged band pattern by southern blotting: a pitfall of genetic diagnosis.

Uemura A, Sugahara K, Nagai H, Murata K, Hasegawa H, Hirakata Y, Tsukasaki K, Yamada

Y, Kamihira S.

Lab Hematol. 2005;11:8–13. doi: 10.1532/LH96.04061.

16)

Identification of biclonal (duplex) leukaemic cells expressing either CD4+/CD8- or CD4-/CD8+ from a patient with adult T-cell leukaemia/lymphoma.

Kondo S, Kotani T, Tsumori S, Narahara S, Aratake Y, Kobayashi M, Takahashi M, Inoue S, Ohtaki S.

Br J Haematol. 1995 Mar;89(3):669-71.

17)

Spontaneous regression in adult T-cell leukemia/lymphoma.

Shimamoto Y, Kikuchi M, Funai N, Suga K, Matsuzaki M, Yamaguchi M.

Cancer. 1993 Aug 1;72(3):735-40.

18)

Lymphoma-type adult T-cell leukaemia-lymphoma with a bulky cutaneous tumour showing multiple human T-lymphotropic virus P1 DNA integration.

Kato N, Sugawara H, Aoyagi S, Mayuzumi M.

Br J Dermatol. 2001 Jun;144(6):1244-8.

19)

Graft-versus-adult T-cell leukemia/lymphoma effect following allogeneic hematopoietic stem cell transplantation.

Yonekura K, Utsunomiya A, Takatsuka Y, Takeuchi S, Tashiro Y, Kanzaki T, Kanekura T.

Bone Marrow Transplant. 2008 Jun;41(12):1029-35.

20)

Poor outcome of autologous stem cell transplantation for adult T-cell leukemia/lymphoma: a case report and review of the literature.

Tsukasaki K, Maeda T, Arimura K, Taguchi J, Fukushima T, Miyazaki Y, Moriuchi Y, Kuriyama K, Yamada Y, Tomonaga M.

Bone Marrow Transplant 1999; 23: 87–89.

21)

High circulating proviral load with oligoclonal expansion of HTLV-1 bearing T cells in HTLV-1

carriers with strongyloidiasis.

Gabet AS, Mortreux F, Talarmin A, Plumelle Y, Leclercq I, Leroy A, Gessain A, Clity E, Joubert M, Wattel E.

Oncogene. 2000;19:4954–60. doi: 10.1038/sj.onc.1203870

22)

Involvement of IL-2/IL-2R system activation by parasite antigen in polyclonal expansion of CD4 (+)25(+) HTLV-1-infected T-cells in human carriers of both HTLV-1 and *S. stercoralis*.

Satoh M, Toma H, Sugahara K, Etoh K, Shiroma Y, Kiyuna S, Takara M, Matsuoka M, Yamaguchi K, Nakada K, Fujita K, Kojima S, Hori E, Tanaka Y, Kamihira S, Sato Y, Watanabe T.

Oncogene. 2002;21(16):2466–75. doi: 10.1038/sj.onc.1205329.

23)

Human T-cell leukemia virus type I (HTLV-I) infection and the onset of adult T-cell leukemia (ATL)

Matsuoka M.

Retrovirology. 2005;26(2):27. doi: 10.1186/1742-4690-2-27.

24)

Clonal expansion of human T-cell leukemia virus type I-infected cells in asymptomatic and symptomatic carriers without malignancy.

Wattel E, Vartanian JP, Pannetier C, Wain-Hobson S.

J Virol. 1995;69(5):2863–8

25)

Two types of defective human T-lymphotropic virus type I provirus in adult T-cell leukemia.

Tamiya S, Matsuoka M, Etoh K, Watanabe T, Kamihira S, Yamaguchi K, Takatsuki K.

Blood. 1996;88(8):3065–3073.

26)

Random integration of HTLV-1 provirus: increasing chromosomal instability.

Ohshima K, Ohgami A, Matsuoka M, Etoh K, Utsunomiya A, Makino T, Ishiguro M, Suzumiya J, Kikuchi M.

Cancer Lett. 1998;132(1):203–212.

27)

Clonal expansion of human T-cell leukemia virus type I-infected cells in asymptomatic and symptomatic carriers without malignancy.

Wattel E, Vartanian JP, Pannetier C, Wain-Hobson S

J Virol. 1995 May;69(5):2863-8.

28)

Quantitative Proviral DNA and Antibody Levels in the Natural History of HTLV-I Infection

Manns A, Miley WJ, Wilks RJ, Morgan OS, Hanchard B, Wharfe G, Cranston B, Maloney E, Welles SL, Blattner WA, Waters D.

J Infect Dis. 1999 Nov;180(5):1487-93.

29)

Rapid quantification of HTLV-I provirus load: detection of monoclonal proliferation of HTLV-I-infected cells among blood donors.

Etoh K, Yamaguchi K, Tokudome S, Watanabe T, Okayama A, Stuver S, Mueller N, Takatsuki K, Matsuoka M.

Int J Cancer. 1999 Jun 11;81(6):859-64.

30)

HTLV-1 provirus load in peripheral blood lymphocytes of HTLV-1 carriers is diminished by green tea drinking.

Sonoda J, Koriyama C, Yamamoto S, Kozako T, Li HC, Lema C, Yashiki S, Fujiyoshi T, Yoshinaga M, Nagata Y, Akiba S, Takezaki T, Yamada K, Sonoda S.

Cancer Sci. 2004 Jul;95(7):596-601.

31)

Human T-cell leukemia virus type I (HTLV-1) proviral load and disease progression in asymptomatic HTLV-1 carriers: a nationwide prospective study in Japan.

Iwanaga M, Watanabe T, Utsunomiya A, Okayama A, Uchimaru K, Koh KR, Ogata M, Kikuchi H, Sagara Y, Uozumi K, Mochizuki M, Tsukasaki K, Saburi Y, Yamamura M, Tanaka J, Moriuchi Y, Hino S, Kamihira S, Yamaguchi K;

Joint Study on Predisposing Factors of ATL Development investigators.

Blood. 2010 Aug 26;116(8):1211-9.

32)

A high-throughput splinkerette-PCR method for the isolation and sequencing of retroviral

insertion sites.

Uren AG, Mikkers H, Kool J, van der Weyden L, Lund AH, Wilson CH, Rance R, Jonkers J, van Lohuizen M, Berns A, Adams DJ.

Nat Protoc. 2009;4(5):789-98. Epub 2009 Apr 30.

33)

Persistent clonal proliferation of human T-lymphotropic virus type I-infected cells in vivo.

Etoh K, Tamiya S, Yamaguchi K, Okayama A, Tsubouchi H, Ideta T, Mueller N, Takatsuki K, Matsuoka M.

Cancer Res. 1997 Nov 1;57(21):4862-7.

34)

Fibroblast-derived induced pluripotent stem cells show no common retroviral vector insertions.

Varas F, Stadtfeld M, de Andres-Aguayo L, Maherali N, di Tullio A, Pantano L, Notredame C, Hochedlinger K, Graf T.

Stem CellP2009 Feb;27(2):300-6.

35)

High-definition mapping of retroviral integration sites identifies active regulatory elements in human multipotent hematopoietic progenitors.

Cattoglio C, Pellin D, Rizzi E, Maruggi G, Corti G, Miselli F, Sartori D, Guffanti A, Di Serio C, Ambrosi A, De Bellis G, Mavilio F.

Blood. 2010 Dec 16;116(25):5507-17.

36)

Identification of HIV integration sites in infected host genomic DNA.

Ciuffi A, Barr SD.

Methods. 2011 Jan;53(1):39-46.

37)

The host genomic environment of the provirus determines the abundance of HTLV-1-infected T-cell clones.

Gillet NA, Malani N, Melamed A, Gormley N, Carter R, Bentley D, Berry C, Bushman FD, Taylor GP, Bangham CR.

Blood. 2011 Mar 17;117(11):3113-22. Epub 2011 Jan 12.

38)

Host sequences flanking the human T-cell leukemia virus type 1 provirus in vivo.

Leclercq I, Mortreux F, Cavrois M, Leroy A, Gessain A, Wain-Hobson S, Wattel E. *J Virol.* 2000 Mar;74(5):2305-12.

39)

Preferential selection of human T-cell leukemia virus type I provirus integration sites in leukemic versus carrier states.

Doi K, Wu X, Taniguchi Y, Yasunaga J, Satou Y, Okayama A, Nosaka K, Matsuoka M. *Blood.* 2005 Aug 1;106(3):1048-53.

40)

Long-term study of indolent adult T-cell leukemia-lymphoma.

Takasaki Y, Iwanaga M, Imaizumi Y, Tawara M, Joh T, Kohno T, Yamada Y, Kamihira S, Ikeda S, Miyazaki Y, Tomonaga M, Tsukasaki K. *Blood.* 2010 Jun 3;115(22):4337-43.

HTLV-1による多段階発癌モデル図

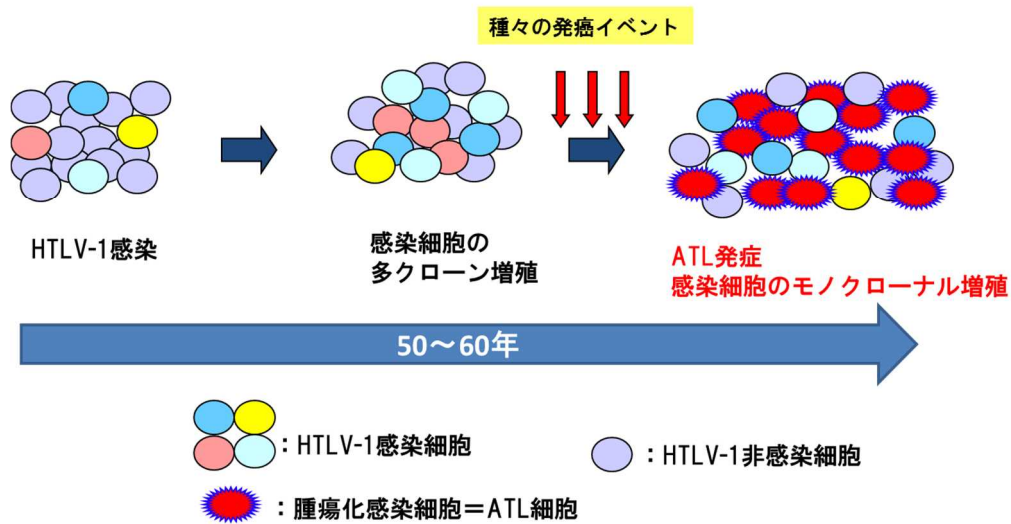


Fig 1 :HTLV-1 による多段階発癌モデル

色の異なる HTLV-1 感染細胞が成人時期までにポリクローナルに増殖し、クローンを構成する。その後様々な腫瘍化イベントを経て 1 クローンのみがモノクローナルに増殖し ATL を発症するというメカニズムを示した。

ATL発症の割合

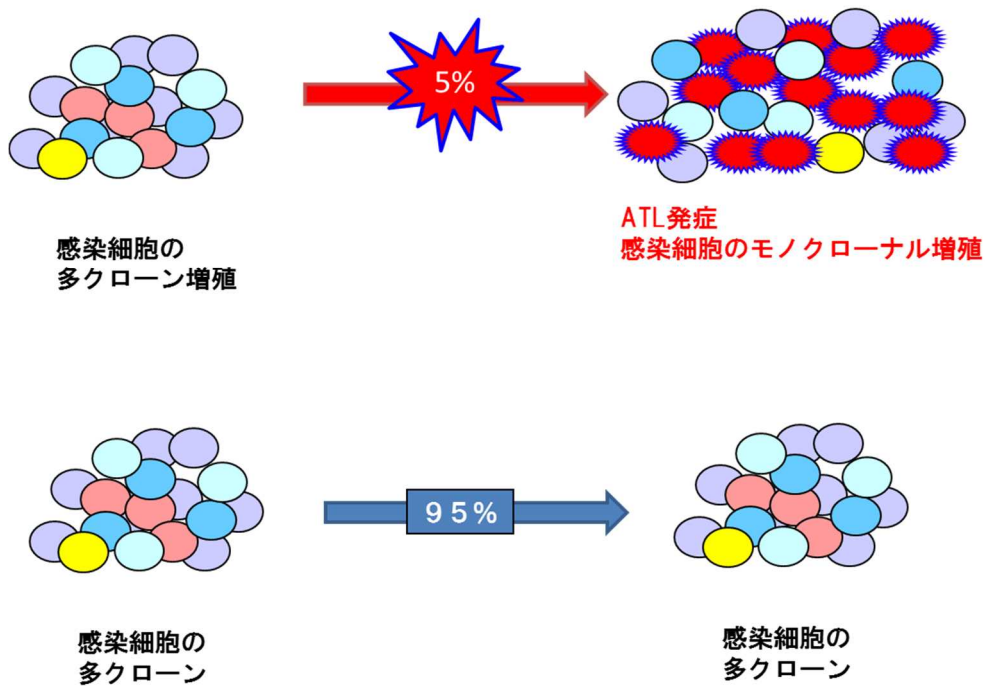
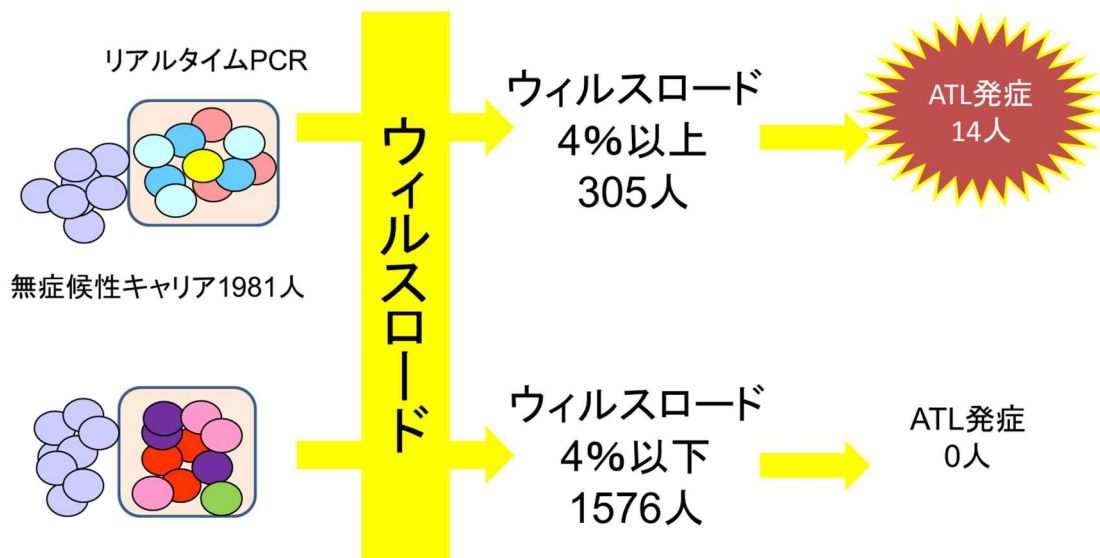


Fig.2 ATL 発症頻度及びクローナリティ
全キャリアの 2~5%が ATL を発症する。95%のキャリアは生涯無症候性のままでポリクローナルな状態のままであるとされている。



Iwanaga.et.al 2010 Blood

Fig.3 ウィルスロードを用いた ATL リスク評価

JSPFAD コホート研究によりウイルスロードの高いキャリアからのみ ATL を発症することが明らかとなった。ウイルスロードは ATL 発症リスク評価のマーカーとなりえることが示唆された。

早期ATLクローン候補検出の条件

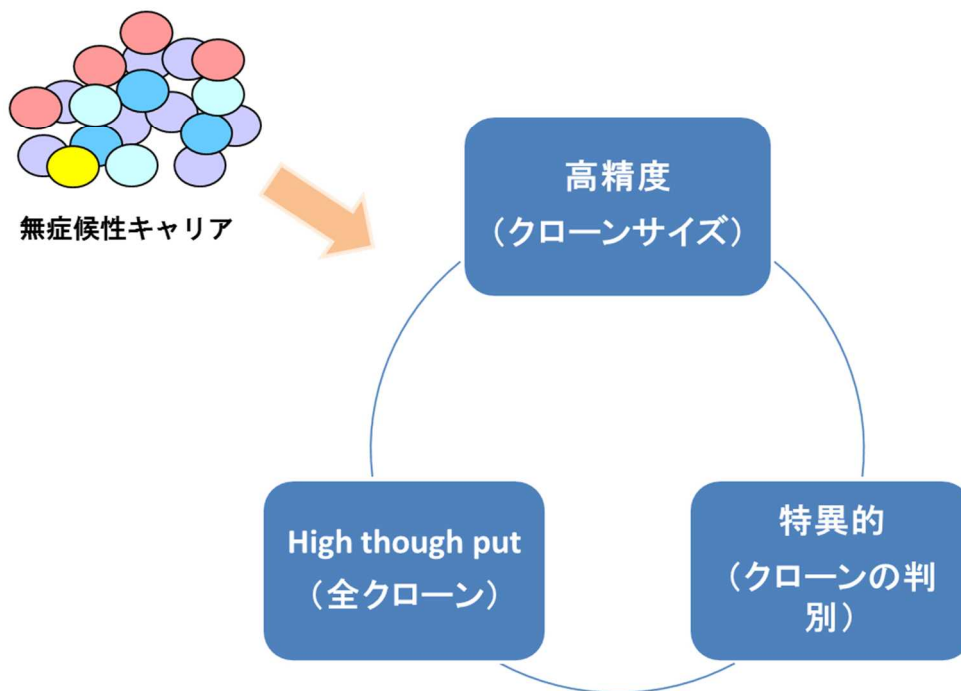
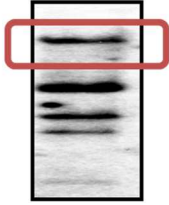


Fig.4 早期 ATL クローン候補検出システムの条件

網羅的にクローンを検出し、各クローンを判別し、そのクローンを構成する感染細胞を正確に測定することが必要である

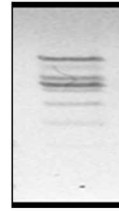
1. サザンブロット解析



メリット
Major cloneの検出が可能

デメリット
感度が低い
high throughputではない

2. Inverse PCR



メリット
高感度(対サザンブロット)

デメリット
再現性が低い
低感度(対Liker mediated PCR)

Fig.5 既存クローナリティ解析システム

代表的なクローナリティ解析システムはサザンブロット解析、Inverse PCR のメリット、デメリットを示した。

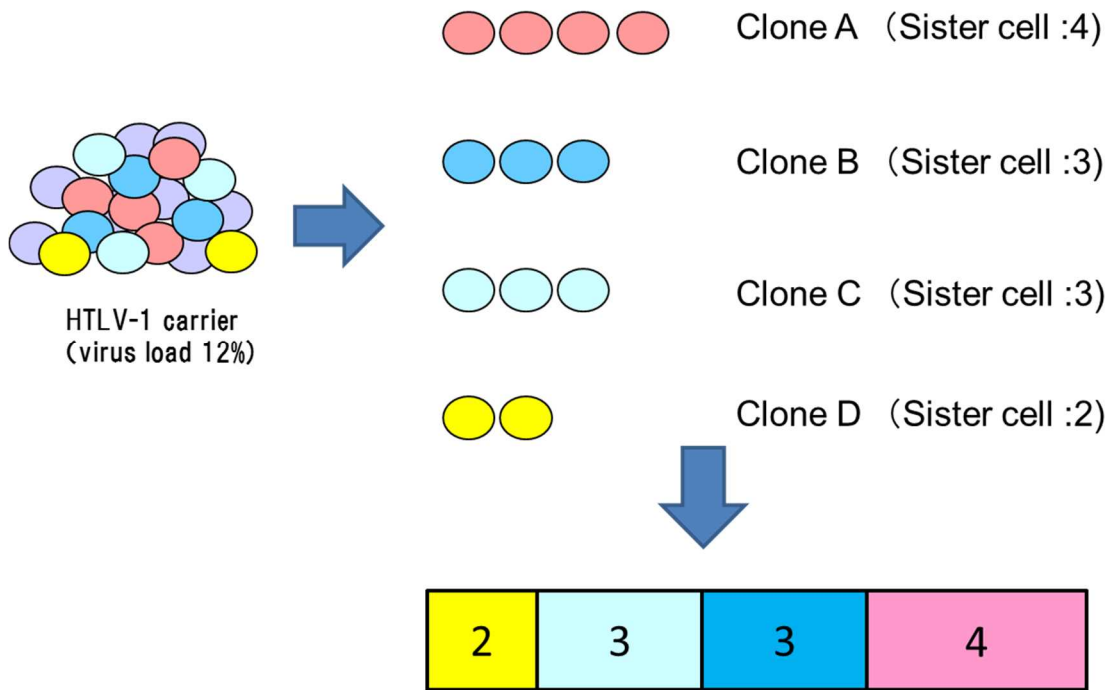


Fig.6 クローナリティ解析の定義

クローナリティ解析とはウイルスロードから得られた全感染細胞の内訳を解析することである。

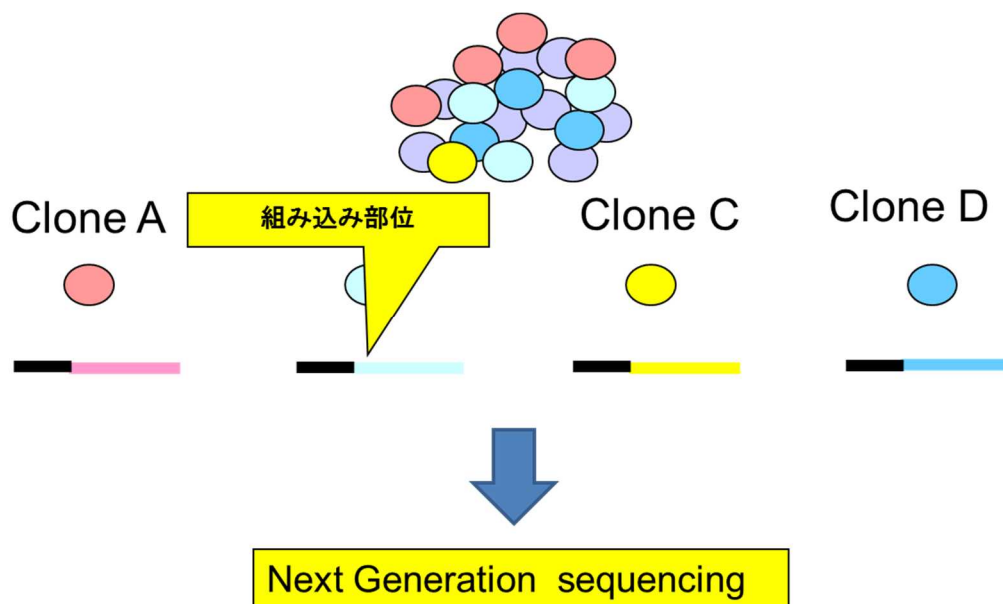


Fig.7 次世代シーケンサーを用いた組み込み部位検出概念図
組み込み部位はクローンにより異なるため、組み込み部位からクローンを判別することができる。

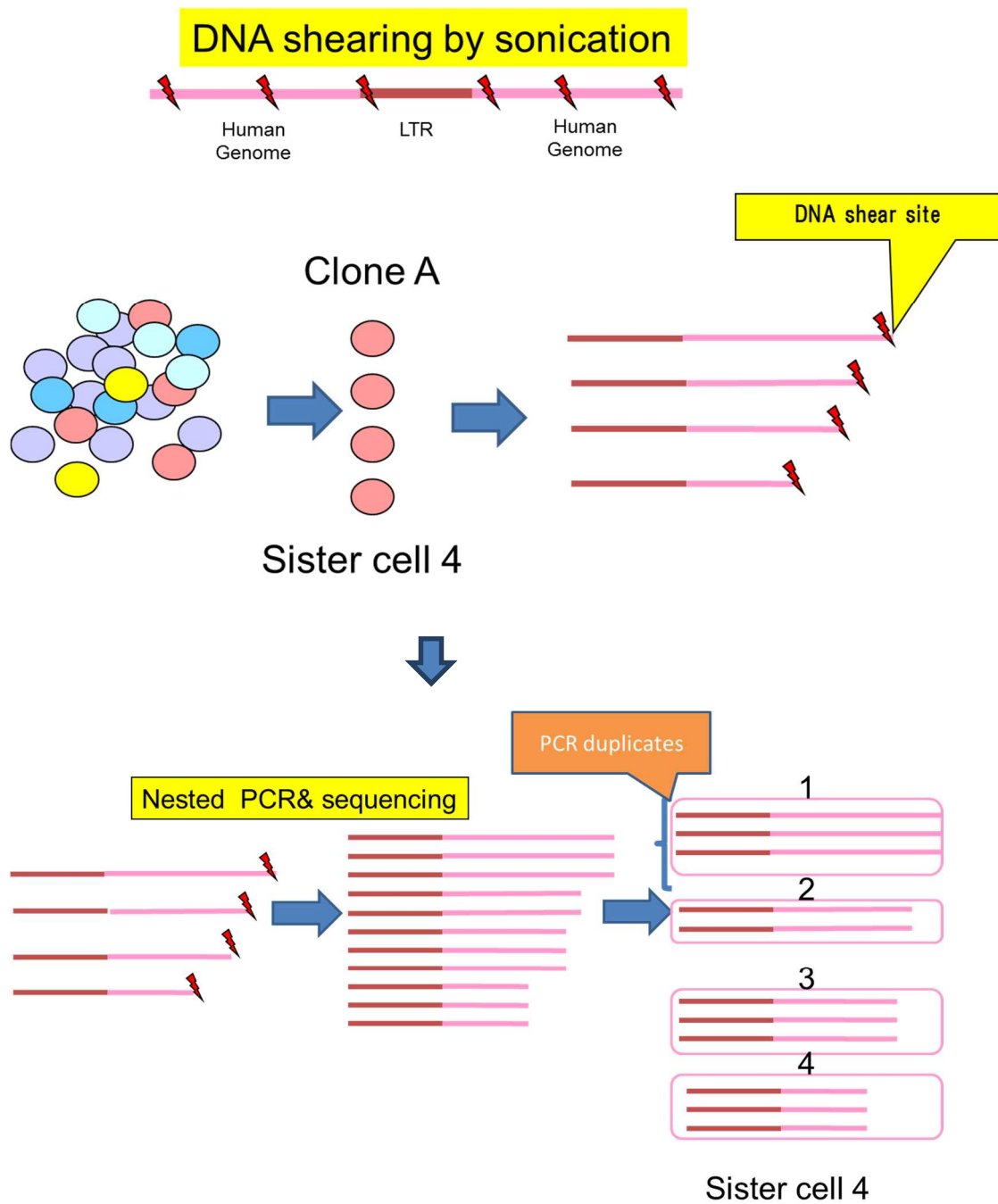


Fig.8 クローンを構成する感染細胞数の計測方法
 シーケンスサンプルを作成する際 sonication 処理により生じる DNA shear site は感染細胞毎に異なる。同一 Shear site をもつ read は PCR duplicate とみなすことで細胞数を計測する。

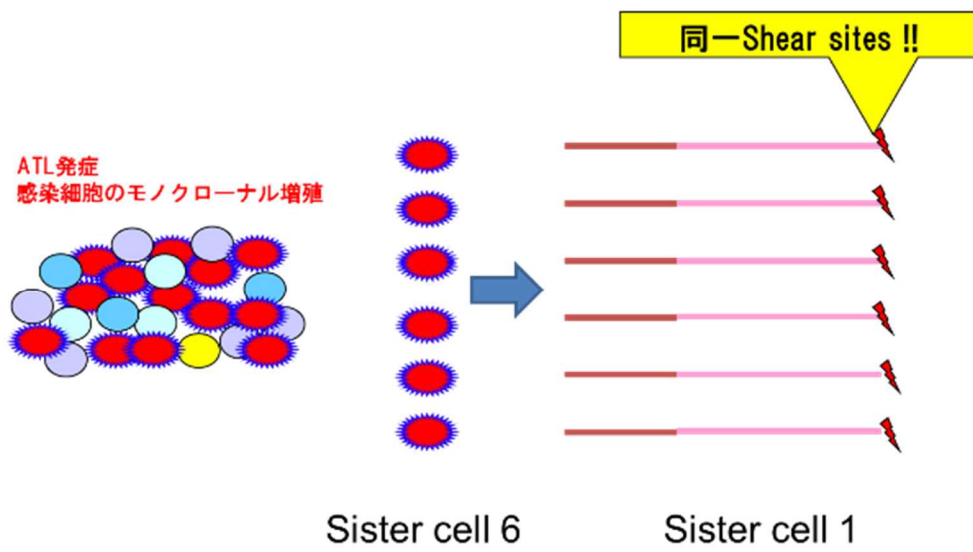


Fig.9 DNA shear site のみを用いた感染細胞計測法の問題点
クローンを構成する感染細胞数が多い場合、別細胞にも関わらず同一 Shear site を生じる可能性がある。

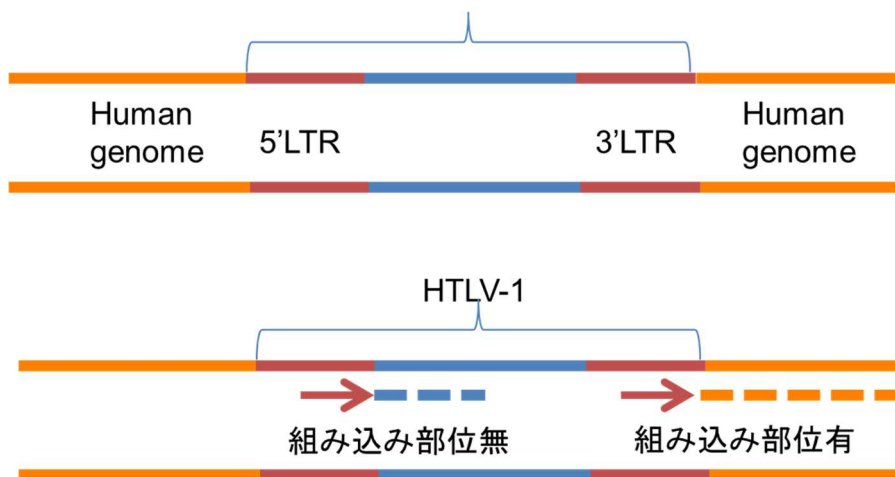


Fig.10 Read1 に含まれる配列情報

LTR 領域は 5'側、3'側に LTR 領域があるため、2 種類の read1 が得られる。
3'側から得られる read1 のみ組み込み部位が含まれている。

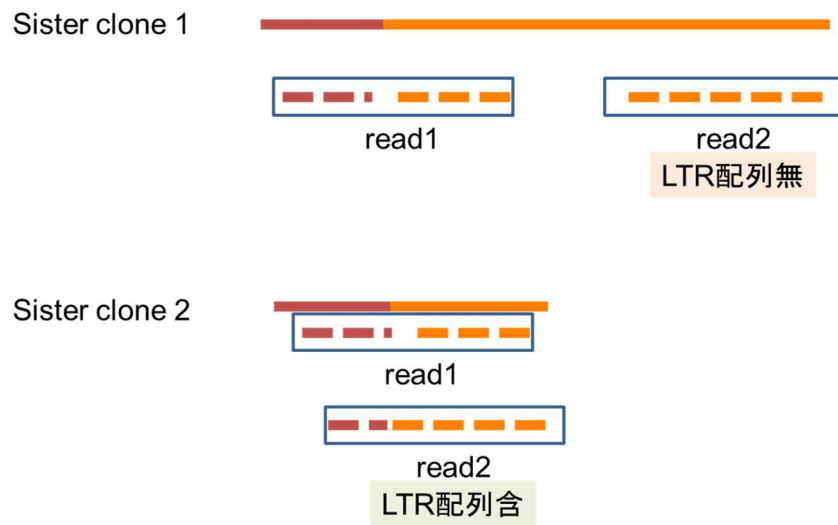


Fig.11 read2 に含まれる配列情報
 LTR 配列の有無により 2 種類の read2 が生じる。

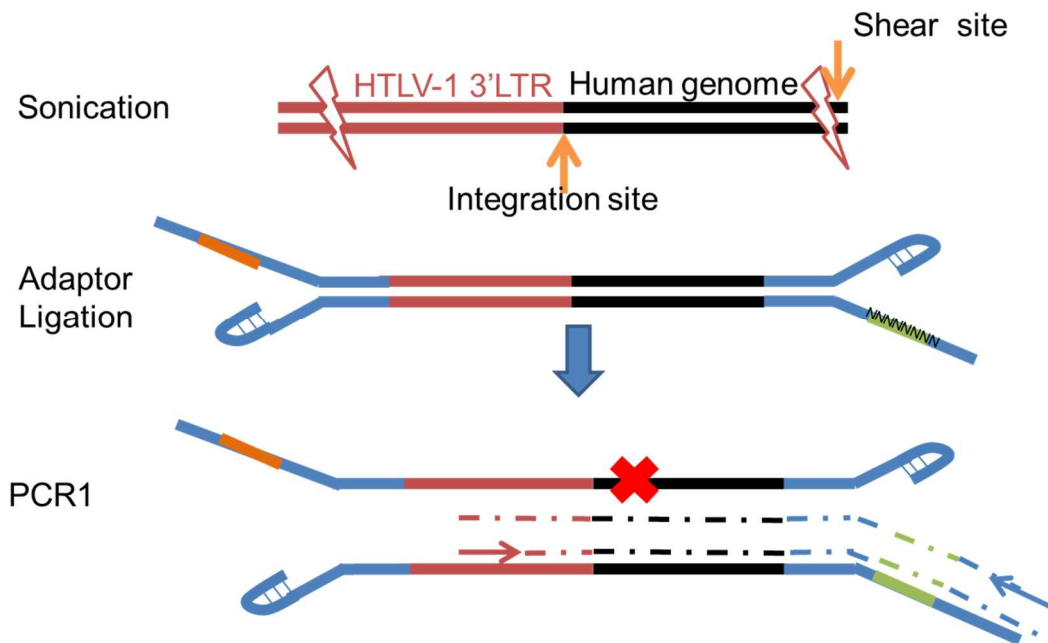
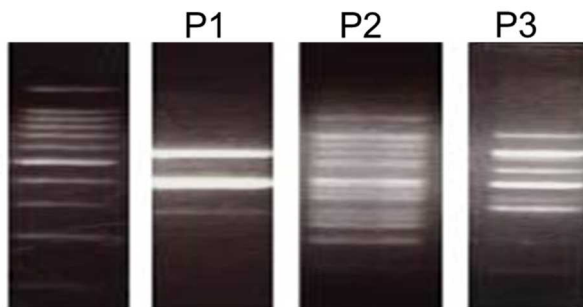


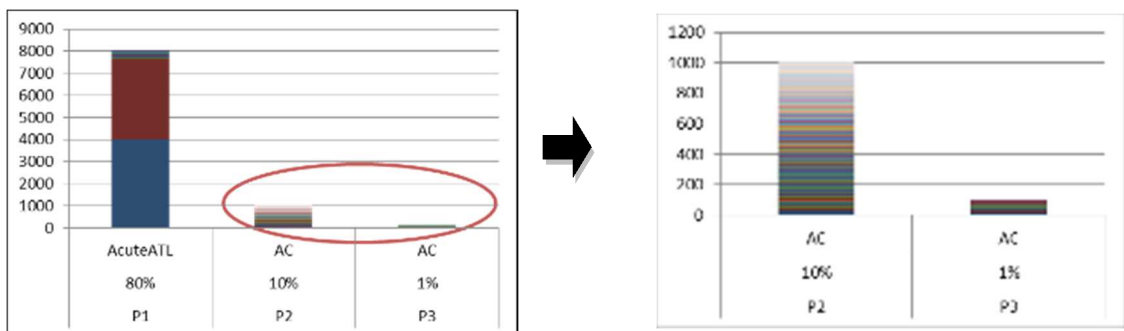
Fig.12 Splinkrette PCR の概念図

3'末端が相補的でないアダプターをライゲーションし、HTLV-1 LTR に特異的なプライマーを用いた **first PCR** を実施する。Reverse プライマーは LTR から増幅された DNA に相補的にデザインされているため、非特異的な増幅を抑えることができる。

A)



B)



C)

sample (VL)	ID	Existing system	Novel system
P1(80%)		3	52
P2(10%)		15	300
P3(1%)		6	53

Fig.13 新規クローナリティ解析システムの検証

A) 既存クローナリティ解析システムを用いた解析(Sau3AIでゲノムを切断後、Splinkrette PCRで増幅)

B)次世代シーケンサーを用いた新規クローナリティ解析 (VLの低い2検体は拡大図を横に掲載)

C)検出されたクローン数の比較

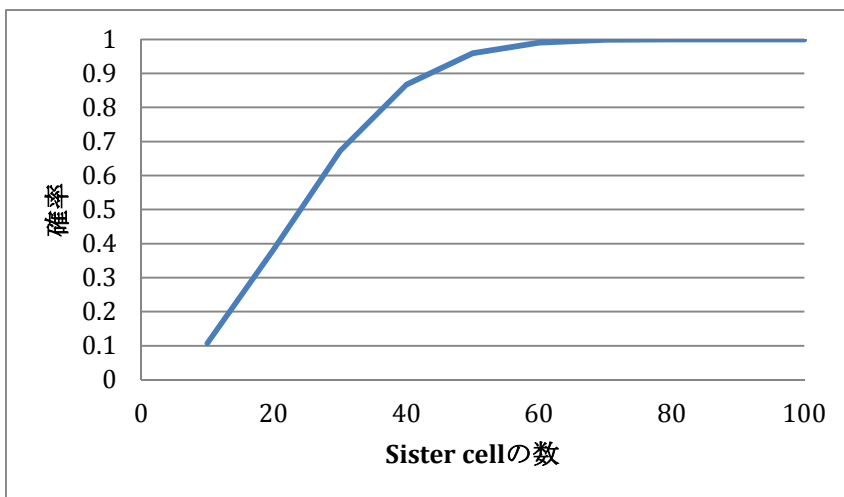
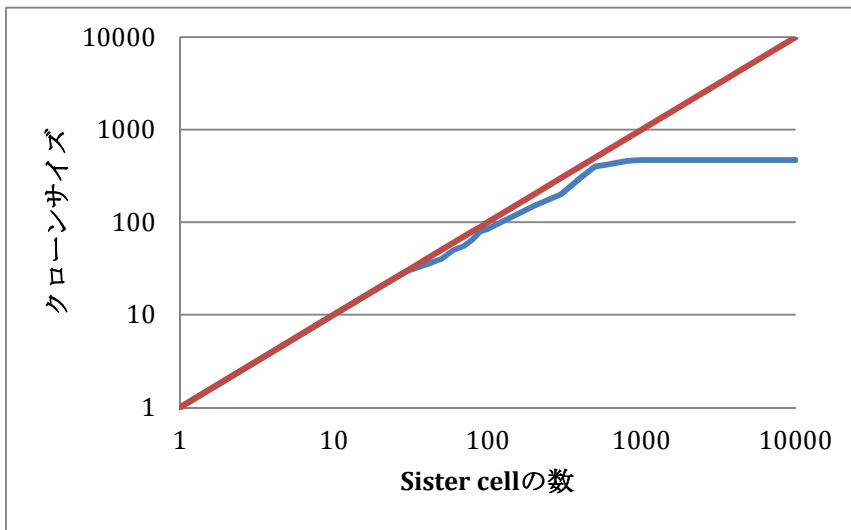


Fig. 14 DNA shear site を用いた場合の問題点

A) Shear site のみ利用した場合と理想のクローンサイズの乖離

青線 : shear site のみ利用した場合

赤線 : タグ配列と shear site を利用した場合

B) 1 クローン中に 2 つ以上の同一 shear site が生じる確率

確率計算式は以下の通り

$1 - 400!/400^n (400-n)!$ n: Sister cell 数

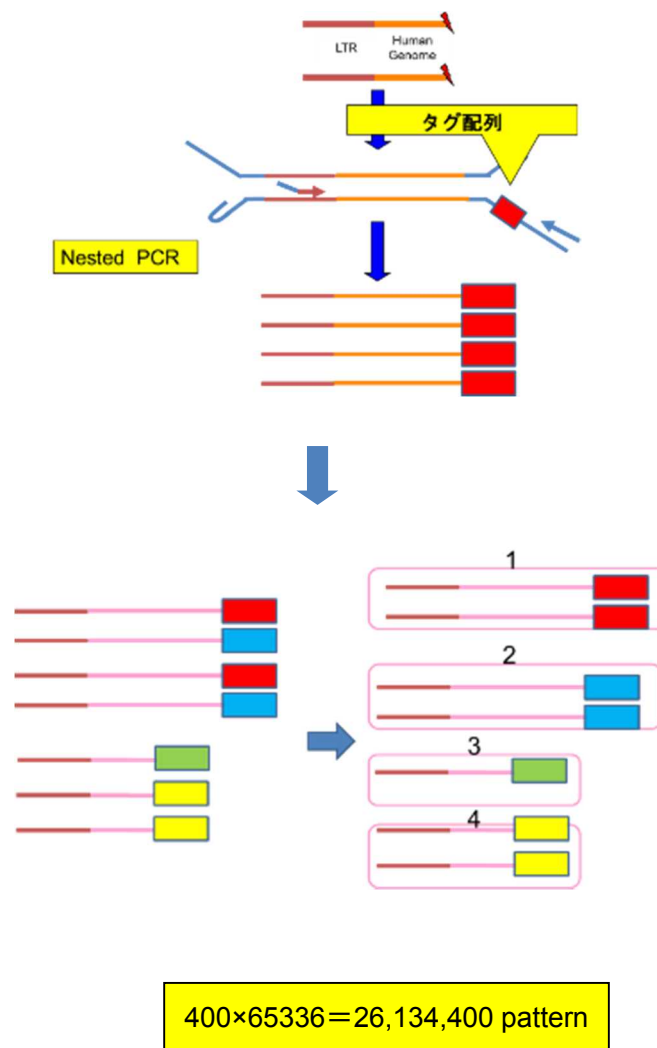


Fig.15 タグ配列を用いた新規クローナリティ解析システムの構築
 8bp のランダムなタグ配列をアダプターに付加し、PCR でゲノムを増幅すると、
 1 感染細胞 HTLV-1 組み込み部位に対して 1 種類のアダプター情報が付加される。
 Shear site とタグ配列を組み合わせて感染細胞数を計測することで全感染細胞
 を網羅することができる。

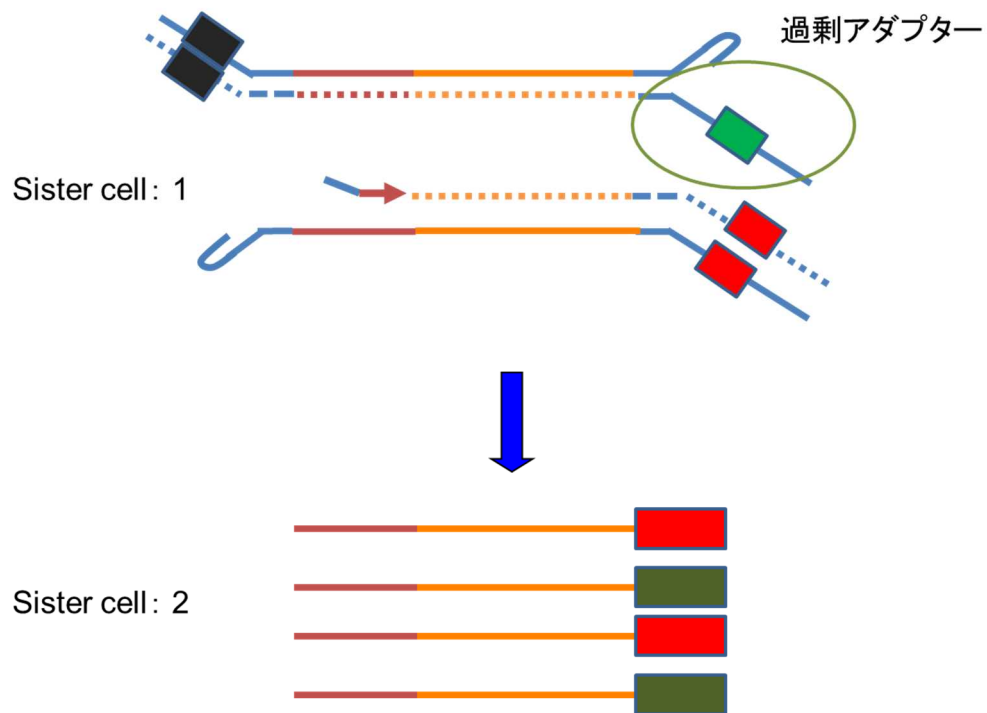


Fig.16 残存アダプターの影響

ライゲーションされなかった過剰アダプターがプライマーとして機能し、非特異的に組み込み部位を増幅する可能性が高い。誤ったクローン構成細胞数を導く恐れがある。

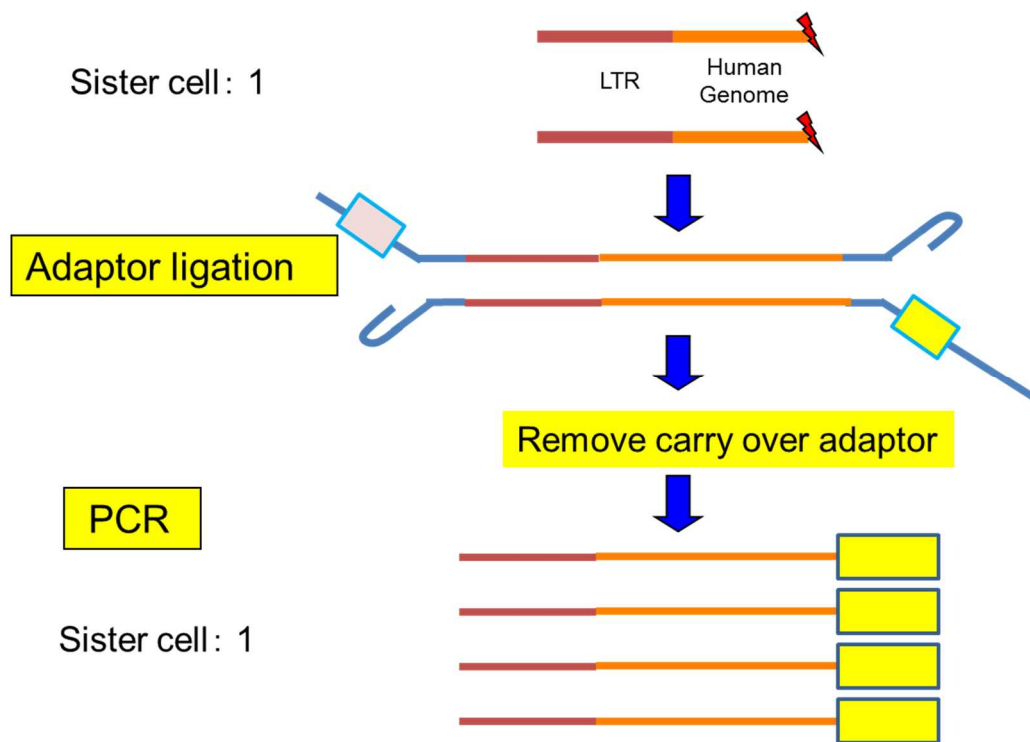


Fig.17 過剰アダプターの除去

アクリルアミドゲルを用いて過剰アダプターを除去することで非特異的な PCR を防ぐことが可能である。

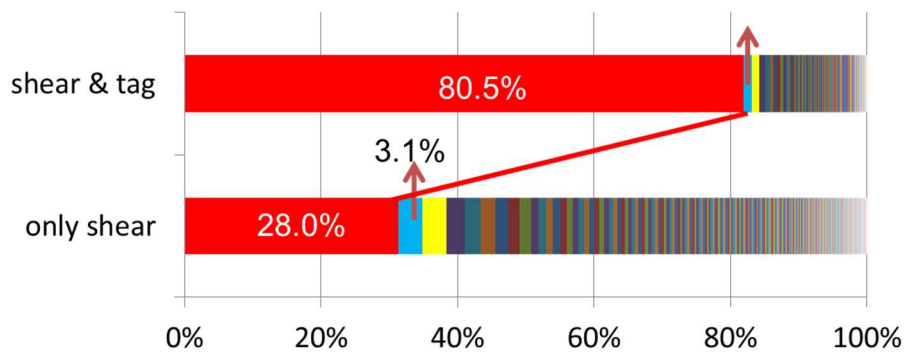


Fig.18 急性型 ATL 検体を用いたタグ配列の有効性の検証
Major clone の割合が大きく異なることが確認された。

	P5(SM)	P6(SM)	P7(CA)
Virus load (%)	30 %	7 %	7 %
M1 (shear)	26 %	22 %	5 %
M1 (shear & tag)	48 %	29 %	11 %
M2 (shear)	19 %	21 %	4 %
M2(shear & tag)	26 %	18 %	5 %
M3(shear)	16 %	16 %	3 %
M3(shear & tag)	11 %	5 %	3 %

Table 1 くすぶり型 ATL、無症候性キャリアの検体を用いたタグ配列の有効性の検証

M1: Major clone

M2: Second major clone

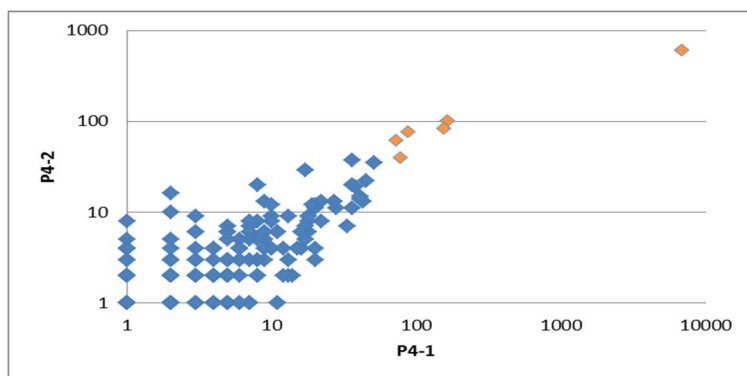
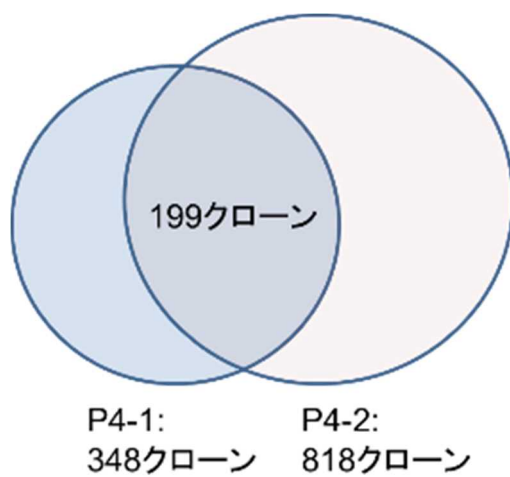
M3: Third major clone

(shear): shear site のみ利用したクローナリティ解析

(shear & tag) タグ配列、shear site 両方を利用したクローナリティ解析

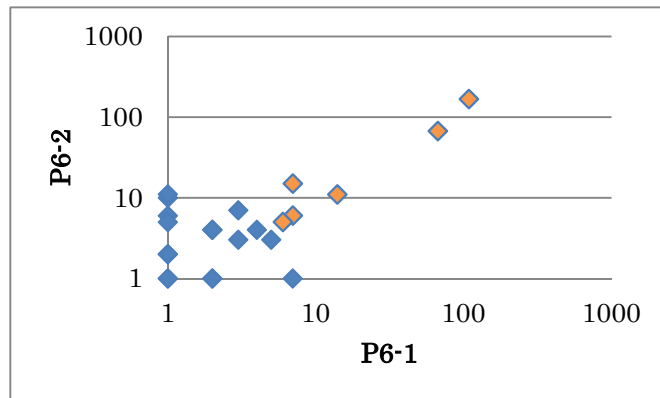
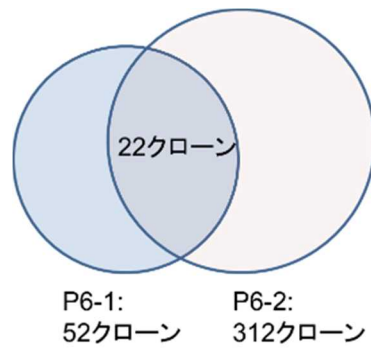
Major clone を持つ ATL 検体はタグ配列によりクローナリティが異なることが確認された。

A)



clone ID	P4-1	P4-2
X-83705335	6768	610
14-30655894	163	102
14-49676330	154	84
6-85461542	87	77
21-35571075	77	40
7-140001924	72	62
16-17339637	51	35
8-120600922	45	22
2-154451370	43	13
20-7574902	41	14
1-192201510	41	15
1-56007282	38	19
13-19529762	36	11
13-33476468	36	37
13-85645211	36	20
12-103725100	33	7
18-19456075	28	11
5-51849120	27	13
16-24733627	22	8
5-78735912	22	13
8-133560618	20	11
3-84669656	20	4
12-122850941	20	3
4-134621319	20	12
6-16274495	19	12
2-158289048	18	8
17-49774114	18	6
4-55715177	18	9
9-66502171	17	5
21-28174788	17	29

B)



Clone ID	P6-1	P6-2
15-59364358	109	168
7-58027425	67	67
Y-9916773	7	15
13-74706144	14	11
7-5795794	3	7
13-53810439	7	6
13-87240687	1	6
13-84177241	6	5
14-89521750	4	4
7-16307378	2	4
3-15911916	2	4
9-33030800	5	3
21-10755510	3	3
13-73872195	1	2
7-86279178	1	2
2-84925282	1	1
17-42095548	2	1
16-13925826	2	1
6-88387461	1	1
3-115225871	7	1
21-442421	1	10
X-107427788	1	5
12-331008	1	11

Fig.19 クローナリティ再現性の検証

A) P4-1,P4-2 検体のクローナリティ解析

Clone ID: HTLV-1 組み込み部位

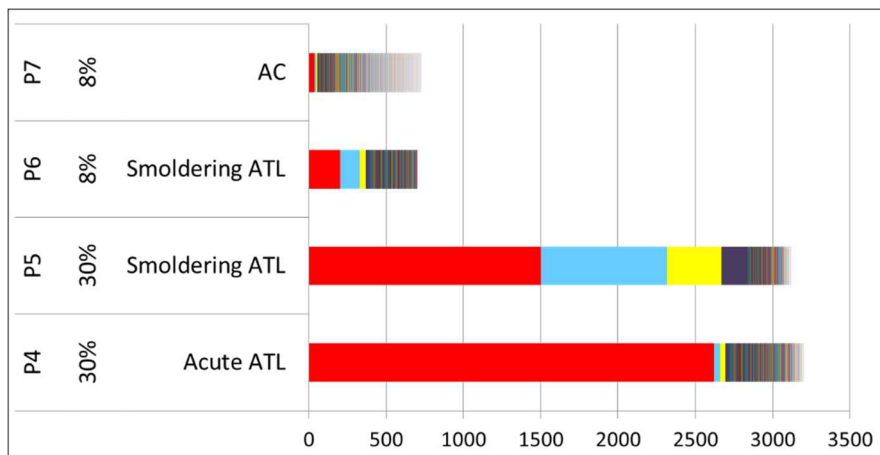
Major clone top10 に含まれるクローンをハイライトした。

B)P6-1,P6-2 検体のクローナリティ解析

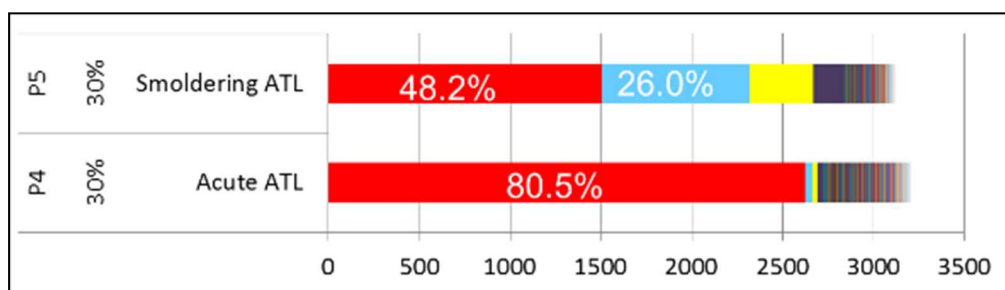
Clone ID: HTLV-1 組み込み部位

Major clone top10 に含まれるクローンをオレンジ色でハイライトした。上位 30 クローンをテーブルに示した。

A)



B)



C)

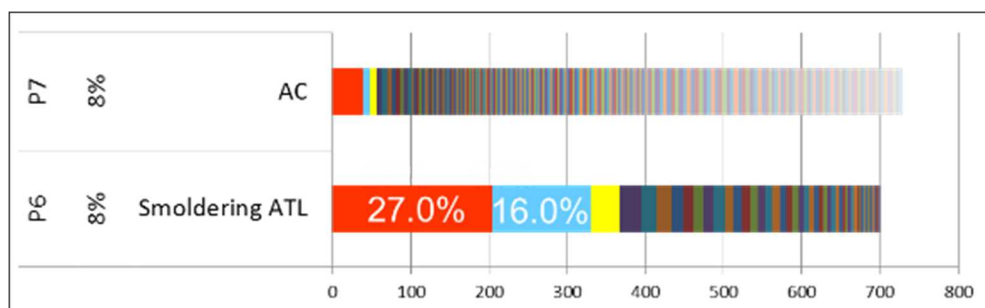


Fig.20 ATL におけるクローナリティ解析

A) ウィルスロードの低い 2 検体及び、ウィルスロードの高い 2 検体のクローナリティ解析

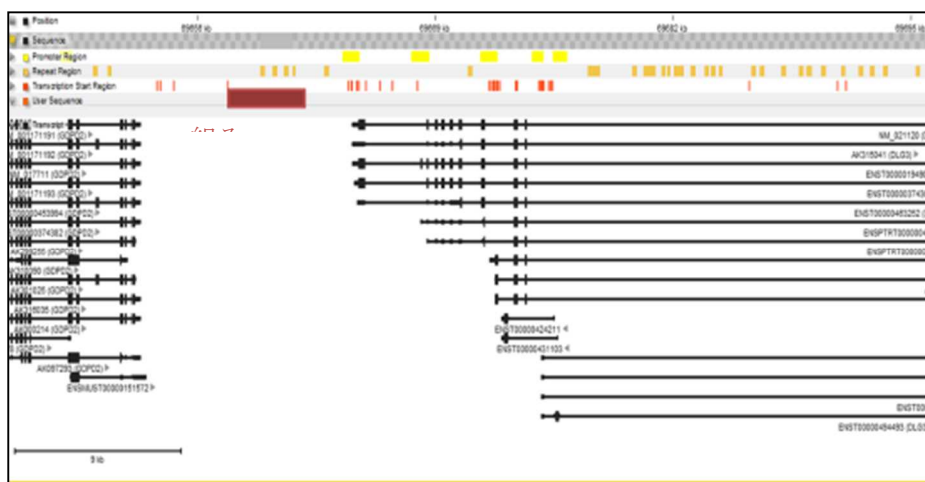
B) ウィルスロードの高い急性型 ATL、及びくすぶり型 ATL のクローナリティの比較

C) ウィルスロードの低いくすぶり型 ATL、及び無症候性キャリアのクローナリティの比較

A)

chromosome	Start	end	ID	strand	Clone size
chrX	69660858	69660958	1	+	4017.118402

Intergenic GDPD2 と DLG3 の間



B)

chromosome	Start	end	ID	strand	Clone size
chr7	34235363	34235463	2	+	3651.92582

BMPER と ENST00000446484 の間

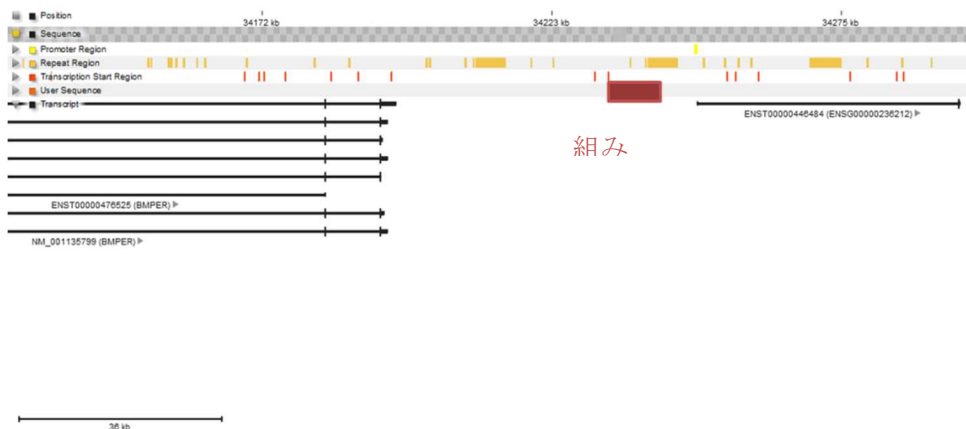


Fig. 21 ATL4 検体の Major clone ヒトゲノム組み込み部位の検証

A) P1 major clone1 の組み込み部位

B) P1second major clone の組み込み部位

C)

chromosome	Start	end	ID	strand	Clone size
chrX	83705335	83705435	1	-	2621.072121

HDX 遺伝子のイントロン

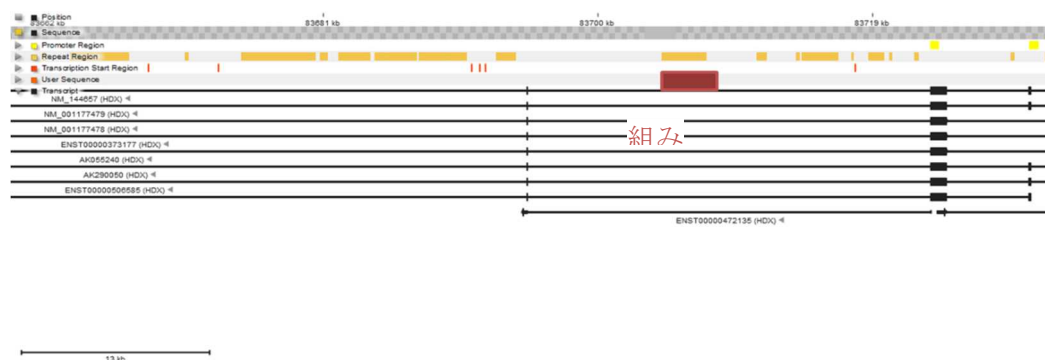


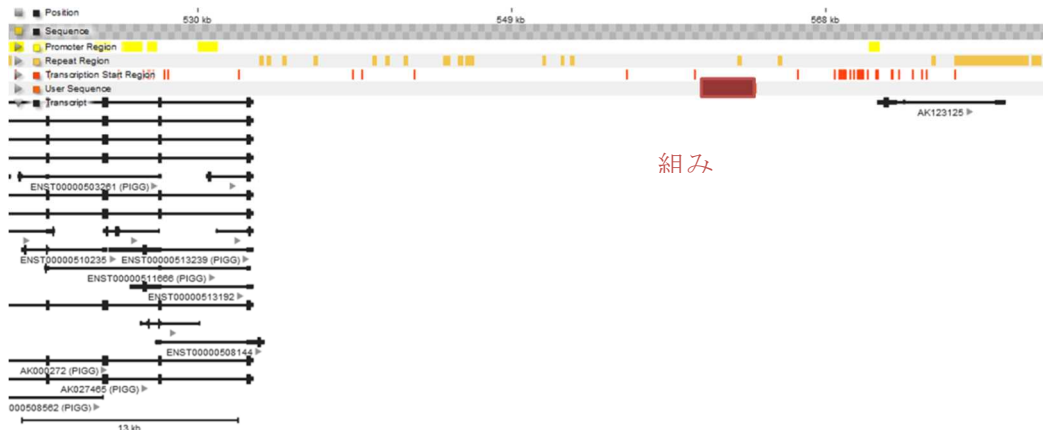
Fig.21 ATL4 検体の Major clone ヒトゲノム組み込み部位の検証

C) P4 major clone の組み込み部位

D) P5(30%) major clone の組み込み部位

chromosome	Start	end	ID	strand	Clone size
chr4	563549	563649	1	-	1503.282391

PIGG 遺伝子と FLJ41130 遺伝子間



E)

chromosome	Start	end	ID	strand	Clone size
chr20	58007370	58007470	2	+	812.4707081

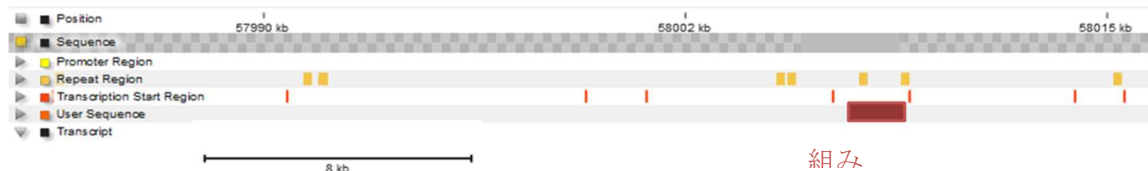


Fig.21 ATL4 検体の Major clone ヒトゲノム組み込み部位の検証

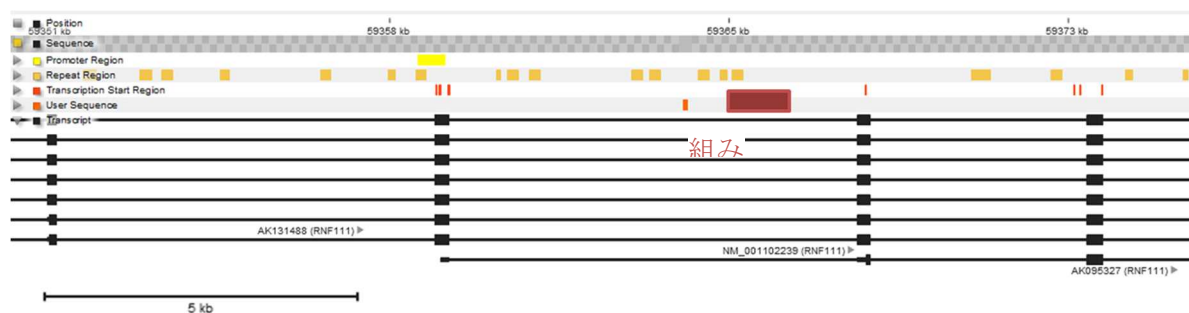
D) P5(30%) major clone の組み込み部位

E) P5(30%) second major clone の組み込み部位

F)

chromosome	Start	end	ID	strand	Clone size
chr15	59364358	59364458	1	+	204.5576

RNF111 のイントロン



G)

chromosome	Start	end	ID	strand	Clone size
chr7	58027425	58027525	2	-	125.7373



Fig.21 ATL4 検体の Major clone ヒトゲノム組み込み部位の検証

F) P6 (7%) major clone の組み込み部位

G) P6 (7%) second major clone の組み込み部位

	sample ID	Type of genomic element	Exonic, complete	Intronic, complete	Intergenic	Promoter
ATL	P6	Number	0	25	97	1
		Percentage	0.00%	20,5%	79,5%	0,8%
		Enrichment Score	0	0,5	1.6	0,3
	P5	Number	7	163	137	9
		Percentage	4,1%	53.10%	50,9%	2.90%
		Enrichment Score	0,6	1.1	0.9	1,1
	P4	Number	13	223	231	16
		Percentage	2,8%	47,8%	49,5%	3.40%
		Enrichment Score	0,7	1,1	1	1,3
	P1	Number	30	560	506	37
		Percentage	2,7%	51,1%	46,2%	3.40%
		Enrichment Score	0,7	1,1	0,9	1,3
carrier	P7	Number	29	577	573	48
		Percentage	2.50%	48.90%	48,6%	4,1%
		Enrichment Score	0,6	1.1	1,0	1,6
	P3	Number	6	52	56	4
		Percentage	5.30%	45.60%	49,1%	3.50%
		Enrichment Score	1,7	0.9	1,0	1,3
	P2	Number	103	1306	1254	112
		Percentage	3.90%	49,0%	47,1%	4,2%
		Enrichment Score	1,0	1,1	0,9	1,6
	Genome	Percentage	4.10%	44.90%	50.90%	2.60%

Table 2 HTLV-1 組み込み部位のゲノムアノテーション解析結果

Number: クローン数

Percentage: 各アノテーションに含まれるクローン数/全クローン数 ×100

Enrichment score: Percentage (各アノテーション)/Percentage(ヒトゲノム)

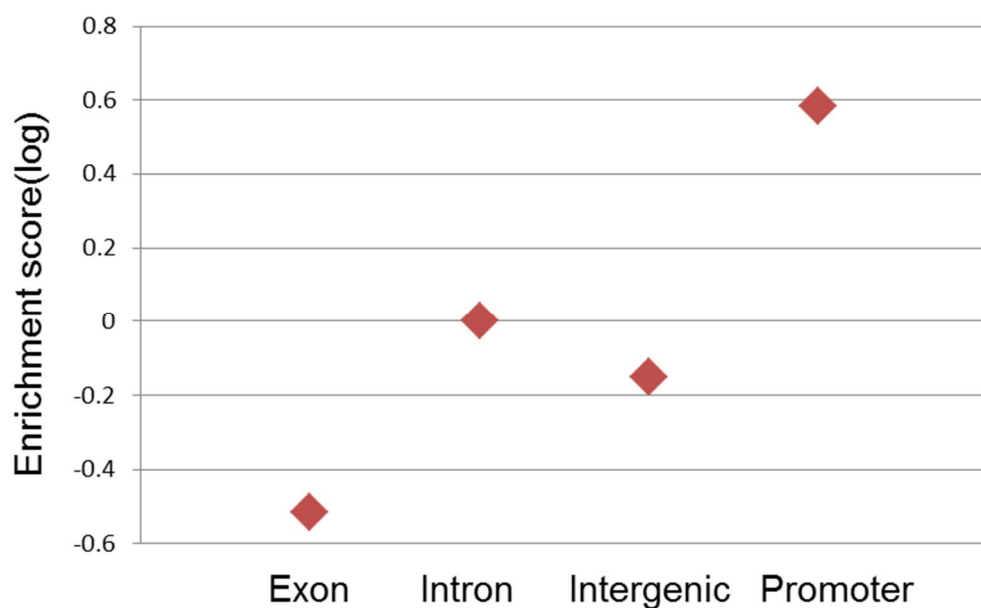


Fig.22 HTLV-1 全組み込み部位のゲノムアノテーションの特徴

MeSH-Term	P-values	Gene(Observed)	Gene(expected)
Joint Diseases	2.32E-03	47	31.81
Speech Disorders	3.80E-03	8	2.51
Lymphatic Diseases	3.99E-03	64	48.11
Leukemia, Prolymphocytic, B-Cell	5.50E-03	2	0.11
Immunoproliferative Disorders	5.79E-03	61	46.02
Pathologic Processes	5.99E-03	129	116.25
Lymphoproliferative Disorders	7.30E-03	60	45.54
Lymphoproliferative Disorders	7.38E-03	59	44.63

Table 3

プロモータ領域が制御する遺伝子群の関連疾患解析結果

赤字：リンパ球関連疾患

Gene(expected):ランダムに選択した場合の遺伝子数

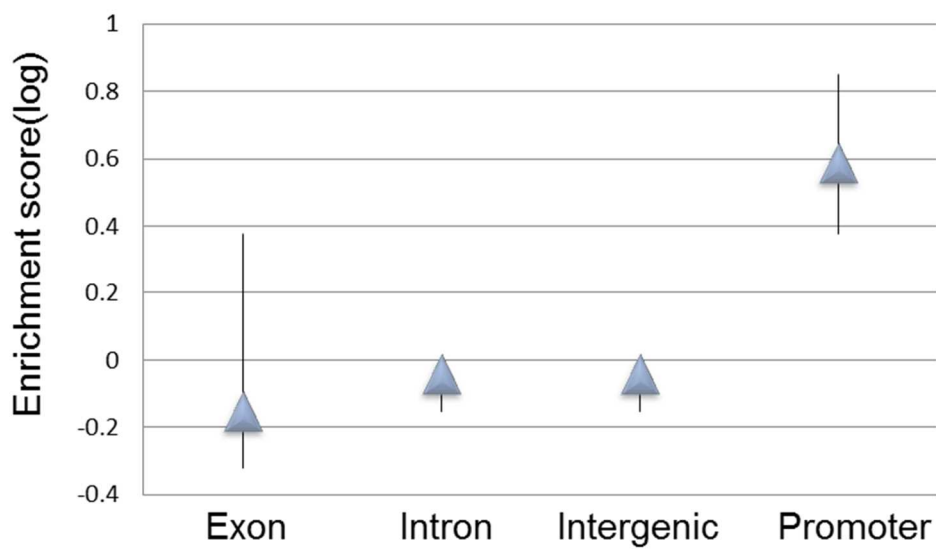
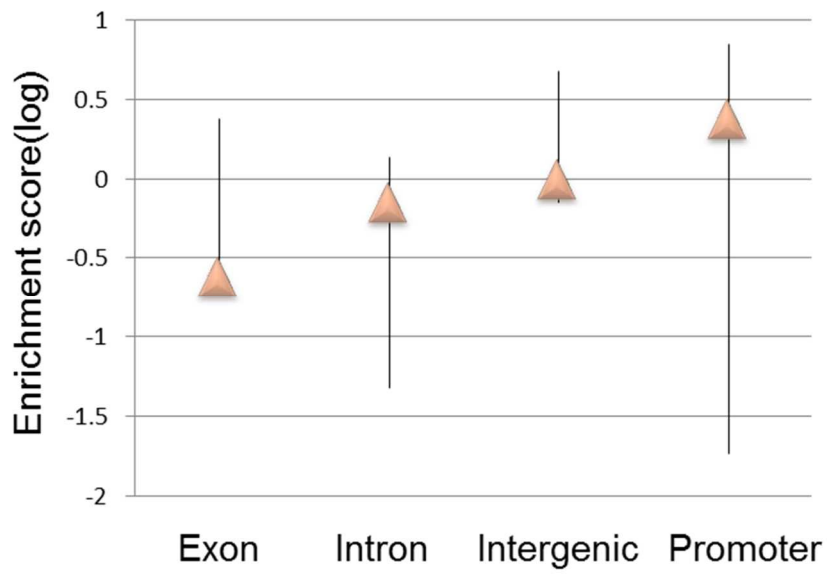
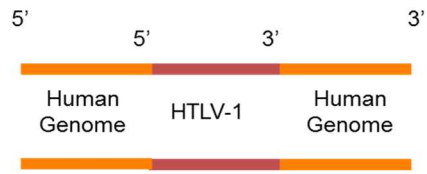


Fig. 23 ATL 検体と無症候性キャリアの HTV-1 組み込み部位の比較

A) ATL4 検体の HTLV-1 組み込み部位特徴解析

B) 無症候性キャリア 3 検体の HTLV-1 組み込み部位特徴解析

ヒトゲノムと同一方向への組み込み



ヒトゲノムと逆方向への組み込み

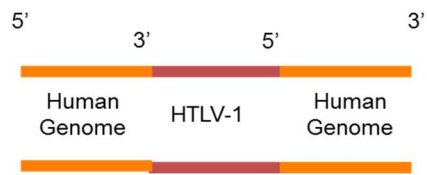


Fig. 24 HTLV-1 組み込み部位の方向性

	sample ID	strand	major clone	middle major clone	minor clone	total clone
ATL	P1 (80%)	+	2	1	11	14
		-	0	2	12	14
	P4 (30%)	+	0	3	105	168
		-	1	1	155	157
	P5 (30%)	+	3	0	102	105
		-	1	0	114	115
P6 (7%)	+	1	2	24	27	
	-	1	1	28	30	

Major clone: 100clone size 以上

Middle major clone: 20 clone size 以上 100 clone size 未満

Minor clone: 20 clone size 未満

	sample ID	strand	major clone	middle major clone	minor clone	total clone
Carrier	P2 (10%)	+	1	8	117	126
		-	1	3	124	128
	P3 (1%)	+	0	1	20	21
		-	0	0	20	20
	P7 (7%)	+	2	1	392	395
		-	1	0	310	311

Major clone: 20 clone size 以上

Middle major clone: 10 clone size 以上 20 clone size 未満

Minor clone: 10 clone size 未満

Table 4 ウィルスゲノム組み込みの方向性とクローナリティの相関検証

Clone size = VL(virus number in 10,000 PBMC)×Sister cell 数/全 HTLV-1 感染細胞数 (各クローンを構成する Sister cell の合計)

Sample ID	Sample type	PVL(/100 PBMC)	total reads	Unique mapping	Number of clone (filter)	Number of clone (no filter)	size selection
S1(ATL_A)	ATL_A	80	50M (100%)	10M (20%)	28	52	なし
S2(AC)	AC	10	30M (100%)	3.7M (10%)	254	300	なし
S3(AC)	AC	1	50M (100%)	5.5M (10%)	41	53	なし
S4-1	ATL_A	30	100M (100%)	22M (22%)	325	466	あり
S4-2	ATL_A	30	100M (100%)	30M (30%)	818	1236	なし
S5	ATL_S	31	97M (100%)	14M (15%)	220	307	あり
S6-1	ATL_S	7	100M (100%)	0.9M (0.9%)	57	122	あり
S6-2	ATL_S	7	100M (100%)	41M (4%)	312	535	なし
S7	AC	7	100M (100%)	15M (15%)	706	1179	あり

ATL_A: 急性型 ATL

ATL_S : くすぶり型 ATL

AC : 無症候性キャリア

Size selection: アクリルアミドゲル電気泳動を用いて、過剰アダプターを除去したサンプルを示す

Table 5 利用検体の内訳

新規クローナリティ解析システムで解析した 7 検体 (9 サンプル) のシーケンスリード数、クローン数を示した。

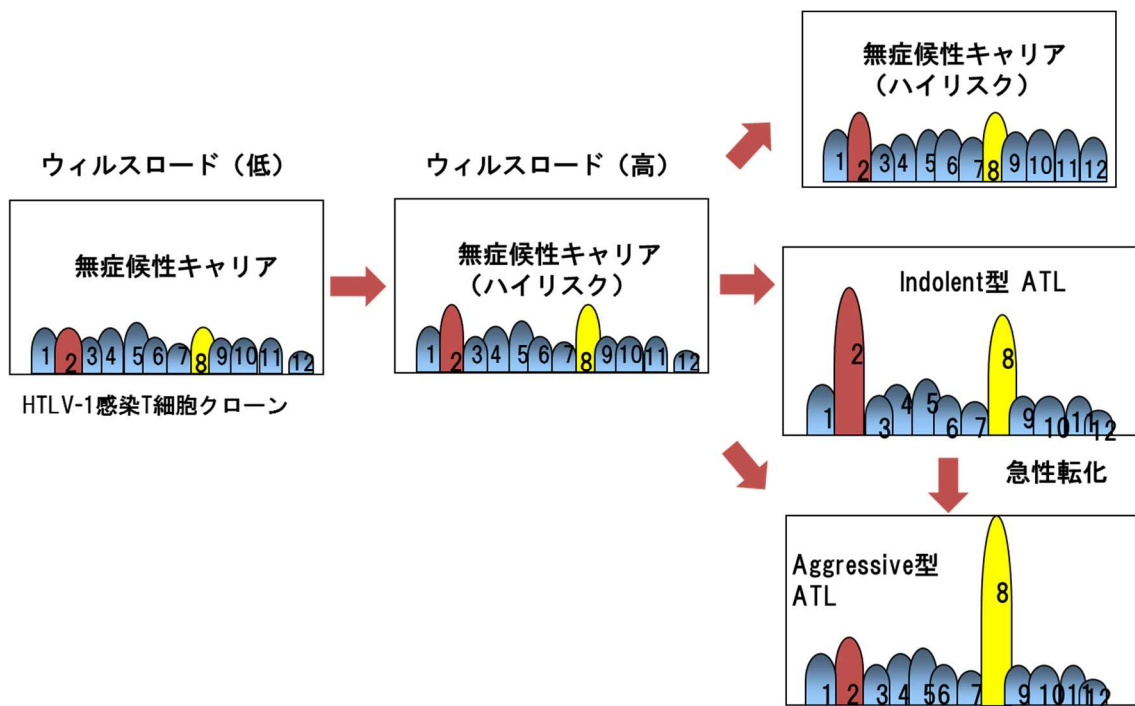


Fig. 25 ATL 発症リスクグループの分類及び、くすぶり型 ATL のクローナリティ概念図

新規クローナリティ解析システムを用いることでウィルスロードでは分類できなかった ATL ハイリスクグループを早期に見つけることが可能となる。また、くすぶり型 ATL は複数の Major clone を持つ可能性が示唆された。