

イベリアトゲイモリにおけるゲノム巨大化と
イントロン領域へのトランスポゾン集積

**Genome Expansion and
Transposon Accumulation in Intronic Regions
of the Iberian Ribbed Newt**

総合研究大学院大学 生命科学研究科 基礎生物学専攻
基礎生物学研究所 再生生物学研究室

保 和人

目次

英語要旨	3-4
日本語要旨	5
第一章：背景と目的	6-9
第二章：結果	
<i>Pw</i> と <i>Xt</i> 間におけるフェリチン重鎖(<i>Fth</i>)遺伝子を含むシンテニー領域のゲノム比較	10-13
<i>Pw</i> の <i>Fth</i> 領域における拡大とトランスポゾンの蓄積	13-17
CAGE/NET-CAGE 解析は TE 挿入により巨大化した遺伝子の転写活性の低下を示す	17-18
<i>Alb</i> 遺伝子座における TE の挿入による遺伝子サイズの肥大化と転写活性の低下	18-21
<i>Vtg</i> 遺伝子座における TE の挿入による遺伝子転写活性の低下と偽遺伝子化	21-23
<i>Vtg</i> 遺伝子クラスター領域における TE による一部 <i>Vtg</i> 遺伝子の破壊	24-25
<i>Pw</i> ゲノムにおける最近転移したレトロトランスポゾンの検出	26-27
<i>Pw</i> 精巢で高発現するユニークな構造を持つレトロトランスポゾンの発見	28-30
第三章：考察	
有尾両生類におけるゲノム巨大化の動的プロセス	31-32
TE 侵入に対する構造的緩衝材としてのイントロン	33-34
なぜイモリでゲノム巨大化が起きたのか？	34-35
転移因子としての <i>Bou1</i> の示唆に富む性質	35-39
遺伝子の subfunctionalization は起きているか？	39-40
イントロンへの TE のランダムな蓄積	40-41
ゲノム巨大化と再生能力の関係	41
第四章：実験材料と方法	42-47
参考文献	48-52
謝辞	53

英語要旨

Vertebrate genome gigantism is frequently accompanied by recurrent bursts of transposable element (TE) activity, yet how such bursts reshape gene architecture and transcription over evolutionary time remains unclear. Urodele amphibians (newts and salamanders) provide an extreme natural experiment: their genomes can exceed 20 Gb while maintaining broadly conserved protein-coding repertoires. I asked how genes tolerate persistent TE invasion without immediate catastrophic disruption of coding sequences and regulatory landscapes.

In this study, I aimed to define how TE accumulation contributes to genome expansion through gene-body inflation and to test whether TE-driven intron expansion is coupled to reduced transcription across loci. I focused on the Iberian ribbed newt *Pleurodeles waltl* (*Pw*; assembled span ~20.3 Gb) and compared it with the compact genome of the western clawed frog *Xenopus tropicalis* (*Xt*; estimated genome size ~1.5 Gb). To capture both long-term accumulation and very recent events, I also compared two independently assembled *Pw* genomes generated from inbred lineages in Japan and Europe. To connect genome structure with transcription, I integrated gene and repeat annotations with high-fidelity (HiFi) genome assemblies and three expression datasets: RNA sequencing (RNA-seq), Cap Analysis of Gene Expression (CAGE; total RNA), and Nascent-Transcript Cap Analysis of Gene Expression (NET-CAGE; nascent RNA). I quantified expression as reads per million (RPM) for each dataset and evaluated locus structure using multi-track genome-browser visualizations and Harr-plot comparisons of syntenic regions between the two *Pw* assemblies. Within gene bodies, I calculated repeat occupancy as the fraction of gene-body bases annotated as TEs and summarized contributions from major TE classes.

I first examined a conserved syntenic region containing a ferritin heavy chain-family (*Fth*) gene cluster and twelve orthologous neighboring genes. In *Xt*, four compact *fth* genes reside within a short interval, whereas in *P. waltl* seven *Fth* paralogs occupy a dramatically expanded region while flanking gene order remains largely conserved. Within this locus, I observed a continuum of intronic TE invasion: some paralogs remain minimally expanded with few intronic insertions, whereas others are enlarged several-fold by diverse TE classes. Importantly, even in strongly inflated paralogs, transcriptome-

supported mRNA sequences preserve an intact open reading frame (ORF), consistent with efficient splicing-mediated removal of TE-containing introns.

Despite frequent ORF preservation, transcriptional profiling revealed that gene-body inflation is commonly associated with reduced transcription. Across *Fth* paralogs, CAGE, NET-CAGE, and RNA-seq signals were inversely related to gene-body length, consistent with transcriptional attenuation for enlarged paralogs.

I then tested whether this relationship generalizes by analyzing two additional liver-enriched loci: albumin (*Alb*) and vitellogenin (*Vtg*). At the *Alb* locus, three conserved albumin-family genes show large TE-driven intron expansion without changes in exon number and open reading frames, while transcription decreases as gene bodies enlarge. The *Vtg* cluster exhibited an even broader spectrum of outcomes, including a massively inflated *Vtg* locus with extensive TE accumulation and markedly reduced transcription. I also observed cases in which transcriptional repression coincides with detectable coding disruption and dispersed exon-like remnants between intact *Vtg* genes, suggesting that prolonged TE attack can eventually fragment gene architecture.

To evaluate whether genome expansion is ongoing, I compared syntenic segments between the Japanese and European *Pw* assemblies and detected lineage-specific insertions, including a recent long terminal repeat (LTR) retrotransposon integration in the downstream region of an *Fth* gene in only one lineage. Additional small insertions and residual fragments of other mobile elements in the same interval further supported very recent activity. These differences between closely related inbred lineages provide direct evidence that TE insertion continues in the new genome and that genome gigantism is an active, ongoing process.

Finally, I identified *Boul*, a testis-enriched, TE-derived locus encoding a reverse transcriptase (RT) and RNase H-like region and, in an alternative reading frame, recombinase-like proteins. Although its molecular role remains hypothetical, its strong testis-biased expression highlights a germline context in which TE-derived genes and TE dynamics may still be shaped by host regulation. Taken together, my results portray genome gigantism as a punctuated yet persistent process driven by successive waves of TE invasion that variably inflate gene bodies, reduce transcription, and diversify gene fates. I propose an evolutionary model in which introns act as protective buffers that absorb TE insertions, helping shield coding exons from immediate disruption and delaying gene decay until transcription is reduced and selective constraints are relaxed.

日本語要旨

ゲノムの巨大化は、長らくトランスポゾン (TE) の反復的な爆発的活性化 (バースト) と関連付けられてきたが、そうしたバーストが進化の過程で遺伝子構造をどのように再構築するのかについては、ほとんど理解されていなかった。本研究では、日本と欧州で独立に構築されたイベリアトゲイモリ (*Pleurodeles waltl*: *Pw*, 20.3Gb) の2つのゲノムアセンブリと、ネッタイツメガエル (*Xenopus tropicalis*: *Xt*, 1.5Gb) のコンパクトなゲノムとを比較することにより、イモリのゲノム拡大期において過去および現在進行形で TE バーストが生じている証拠を見出した。多コピー遺伝子の解析から、イントロンへの TE 侵入の連続的な波によって引き起こされる遺伝子領域 (gene body) の拡大過程が明らかになった。一部の遺伝子は、TE 挿入による著しい遺伝子領域の肥大化 (inflation) にもかかわらず、スプライシングによる TE 配列の効率的な除去を通じて機能を維持していた。一方で、転写抑制を受け、その後変異の蓄積、偽遺伝子化、そしてエクソン-イントロン構造の進行的な崩壊 (burst) へと至る遺伝子も存在した。2つの *Pw* ゲノムアセンブリ間の直接比較は、TE 挿入が現在も生じていることを示しており、ゲノム拡大が能動的かつ現在進行形のプロセスであることを示唆している。さらに我々は、生殖系列に関連した TE 活性と一致して発現する、精巢特異的な TE 由来遺伝子座 *Boul* を同定した。これらの結果は、ゲノム巨大化が断続的かつ持続的なプロセスであることを描き出すとともに、イントロンが TE の侵入に対する防御緩衝材 (protective buffer) として機能し、選択圧が緩和されるまで遺伝子の崩壊を遅らせる機能を有することを示唆している。

第一章 背景と目的

有尾両生類は脊椎動物の中でも突出した再生能力を示している (Agata & Inoue, 2012)。とりわけイモリ科は長年にわたり再生研究の主要モデルとして用いられてきた。これまでに四肢 (Suzuki et al., 2024)、関節 (Tsutsumi et al., 2015)、さらには脳の一部 (Urata et al., 2018)、心臓 (Uemasu et al., 2022)、顎 (Kurosaka et al., 2008)、眼 (Inoue et al., 2012; Maki et al., 2007; Mizuno et al., 2002)、など多様な器官や組織を対象に研究成果が蓄積されている (図 1)。

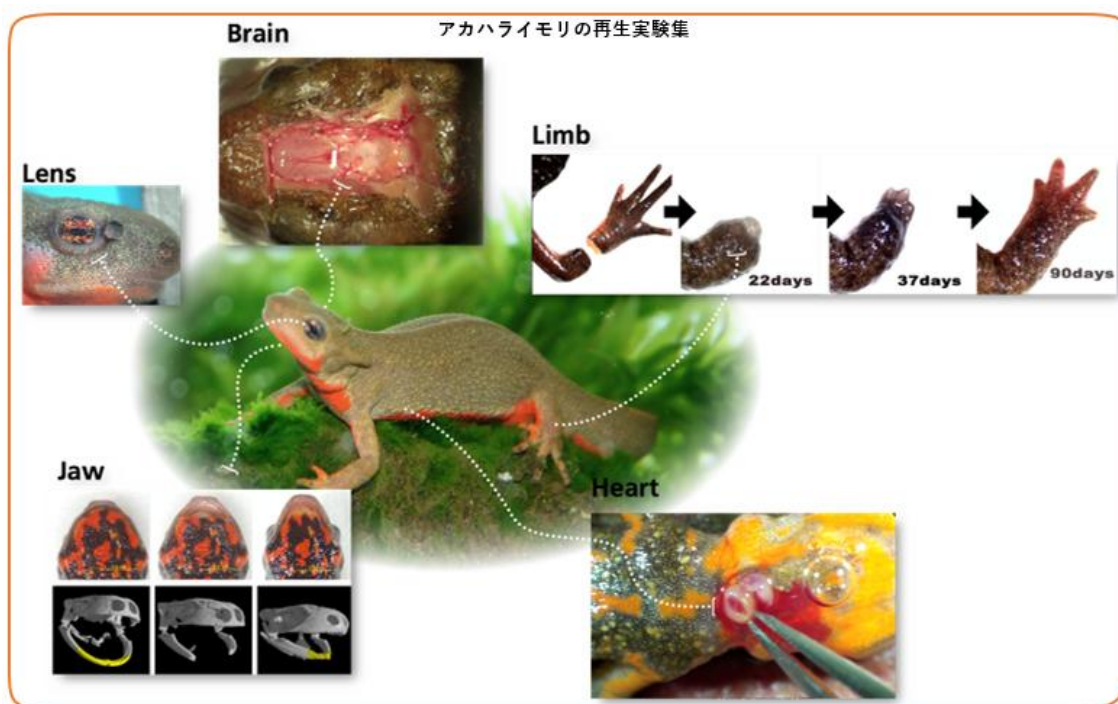


図 1 多様な組織や器官を再生できるイモリ

成体のイモリを用いての様々な実験により有尾両生類の高い再生能が示されてきた。ここでは、眼の水晶体と網膜・脳・四肢・下顎・心臓の一部の再生を例としてとりあげている。その他にも尾も脊髄を含んで再生することが知られている。

一方、同じ両生類の無尾両生類（カエル）では幼生期は肢や脳が再生することが可能であるが、変態後には再生能が低下し、後肢切断後には指を伴わないスパイク状組織の伸長（スパイク再生）が観察される (Yokoyama et al., 2011; Tsutsumi et al., 2016) (図

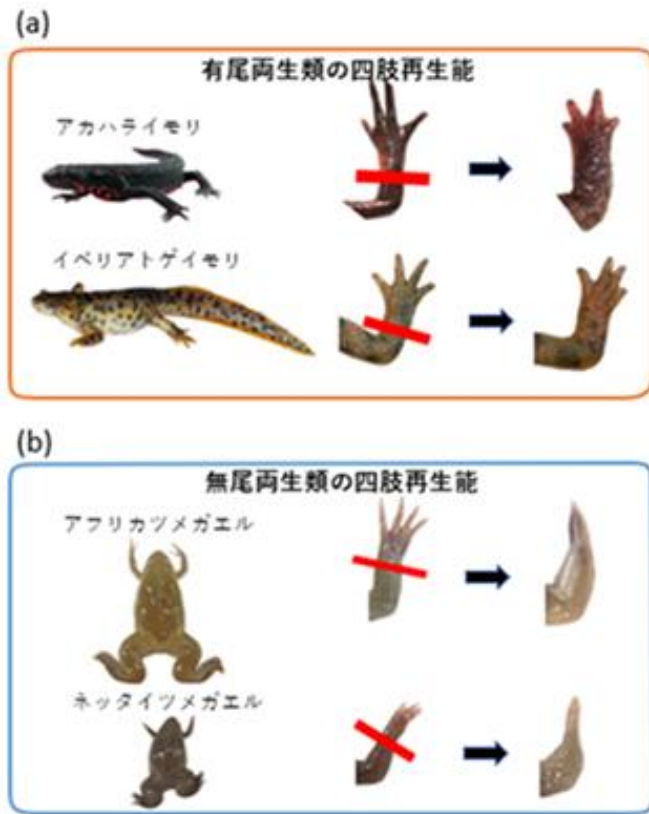


図2 イモリ,カエル,マウスの四肢再生能の違い

(a)有尾両生類は指を伴った四肢を再生できる
 (b)無尾両生類は指の再生を伴わないスパイク状組織が伸長する

2b)。また、哺乳類では軽微な損傷 (Ashley et al., 2012)や

骨折を修復できるものの四肢切断のような大きな欠損は再生せず、イモリが持つような高い再生能力は進化の過程で低下・消失したように見える。このような再生能の違いがどのようなことに起因しているかは大きな謎である。ここでは、両生類における再生能力の違いは、両生類におけるゲノムサイズの顕著な差とも一致することが知られているので、その点に焦点を当てて解析を行った。すなわち、無尾両生類（カエル）のゲノムは一般に数 Gb 規模であるのに対し、有尾両生類は数十 Gb に達する巨大ゲノムを有することが知られてい

る。同じ両生類でありながら再生能の違うイモリとカエルではゲノムサイズに乖離が見られることは、ゲノムの巨大化と有尾類の並外れた再生能力の間に関連があるのかという問いを長らく提起してきた (Elewa, 2017; Sessions & Wake, 2020)。近年の比較ゲノム解析は、当初に考えられていたような有尾両生類の巨大ゲノムは発生時と再生時で異なる遺伝子座にある「遺伝子セットを持つことで巨大になっていたわけではないことを明らかにすると同時に、レトロトランスポゾンや DNA トランスポゾンやなどの転移因子 (TE) が進化的時間を通じて執拗に蓄積した結果として形成されてきたことが示してきた (図 3; Nowoshilow et al., 2018; Schloissnig et al., 2021; Brown et al., 2025; Kimura et al., 2025)。

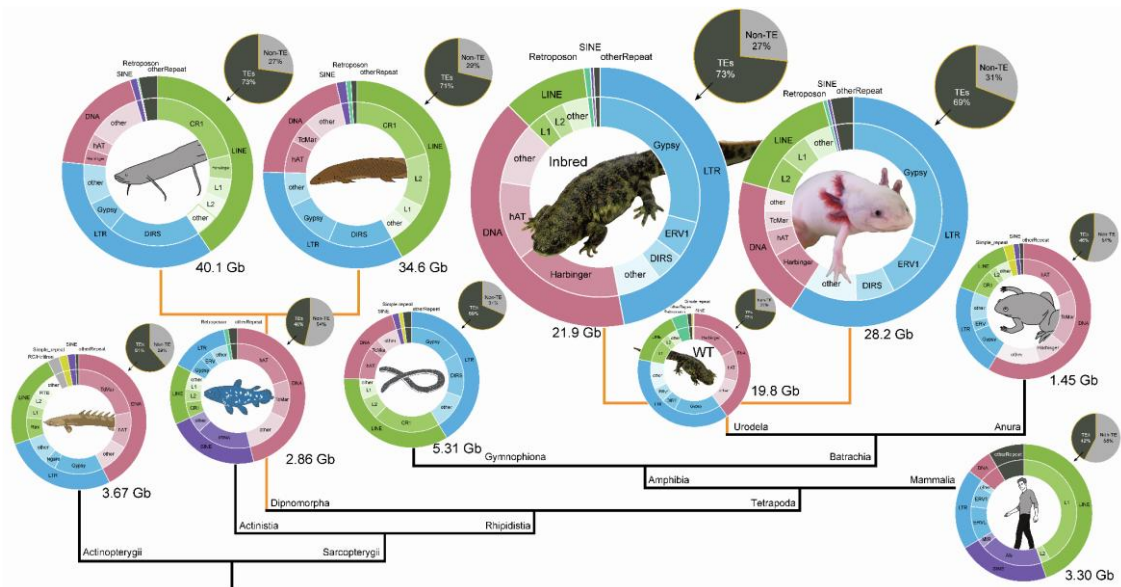


図 3 生物間のゲノムサイズとトランスポゾン(TE)の割合

右下にヒトのゲノム(3.3Gb)と TE の割合が緑グラフで示されている。脊椎動物の系統では、肺魚と有尾両生類(イモリとサンショウウオ)で独立にゲノムの巨大化が進行したと考えられている。ここでは、LTR 型レトロトランスポゾン(水色)、DNA トランスポゾン(赤紫)、LINE(草緑)、SINE(青紫)がゲノム中に占める割合が円グラフで示されている。(Kimura et al., 2025 より転載)

したがって、ゲノムサイズの拡大が再生能力と関係するか否か、またあるとすれば、どのように関係するのかを議論するには、まず有尾両生類におけるゲノム巨大化の分子プロセスを解明することが不可欠である。そこで本研究では、イベリアトゲイモリ(以下 P_w とする (約 20.3 Gb) と ネットアイツメガエル(以下 X_t とする) (約 1.5 Gb, Mitros et al., 2019) のゲノム比較を通じて、 P_w ゲノムが進化の過程でどのようにして巨大化が成し遂げられたのかを推測することを目的とした。

日本では古くからアカハライモリ (*Cynops pyrrhogaster*) がイモリの再生研究のモデルとして使われてきたが、近年になって林らによってイベリアトゲイモリの近交系が確立され、アカハライモリに代わってイベリアトゲイモリが分子生物学的操作を伴う再生研究モデルとして整備され、全ゲノムの解読も進んだ (Hayashi et al., 2013; Kimura et al., 2025)。これに基づき、ゲノム編集システムも次々と作出されている (Suzuki et al., 2018, 2024; Takeuchi et al., 2022)。それに呼応して欧州でもイベリアトゲイモリを用いたゲノム・再生研究が精力的に進められるようになり (Elewa et al., 2017)、小分子 RNA がイモリの脱分化過程に関与する可能性も示唆されている (Subramanian et al., 2023)。さらに近年の高精度 (HiFi) シーケンシング技術の普及により、ゲノムレベルで再生能力を理解する時代を迎えた (Brown et al., 2025; Kimura et al., 2025)。

第二章 結果

Pw と *Xt* 間におけるフェリチン重鎖(*Fth*)遺伝子を含むシンテニー領域のゲノム比較

肝臓において活発に転写されている遺伝子を定量するため、NET-CAGE を実施した。NET-CAGE データの信頼性を評価するため、フェリチン (Ferritin heavy chain)、アルブミン (Albumin)、ビテロジェニン (Vitellogenin) を含む代表的な肝臓高発現遺伝子を解析し、NET-CAGE による発現量解析を従来の CAGE データおよびトランスクリプトームデータと比較した。その結果、これら 3 つの手法のデータセット間で高い相関が認められた。よってこれら 3 遺伝子のゲノム領域に着目し、カエルゲノムとの比較ゲノム解析を行い、イモリ とカエルのゲノム間で非常によく保存されたシンテニー領域 (遺伝子連鎖領域) を同定した(図 4)。

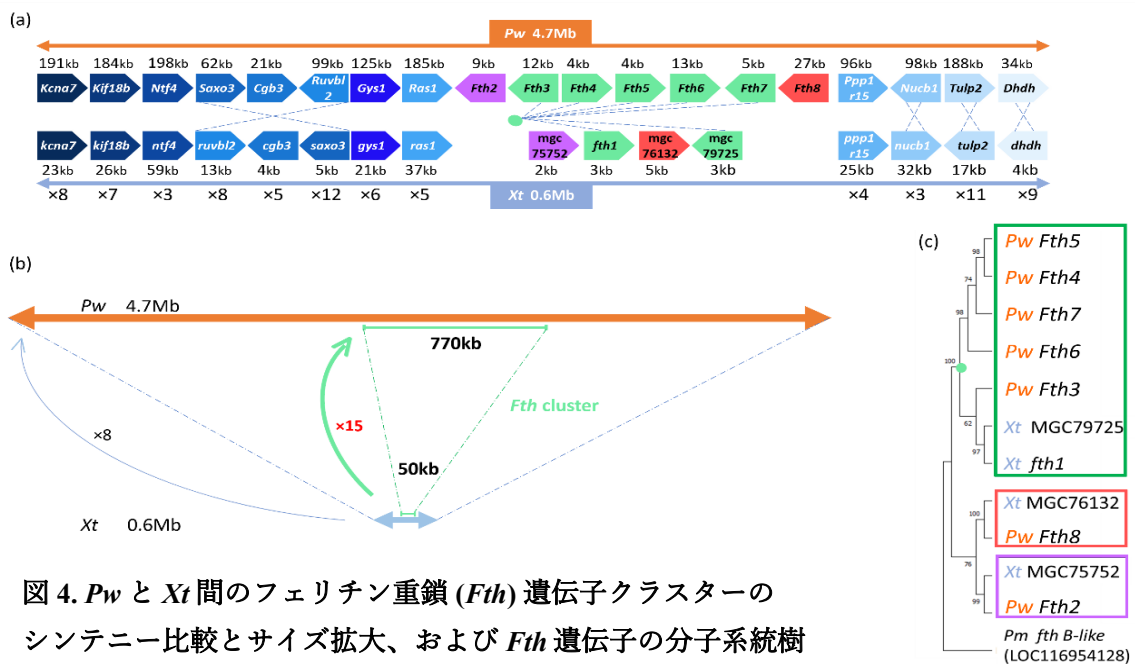


図 4. *Pw* と *Xt* 間のフェリチン重鎖 (*Fth*) 遺伝子クラスターのシンテニー比較とサイズ拡大、および *Fth* 遺伝子の分子系統樹

(a) *Pw* と *Xt* における *Fth* クラスターを含むシンテニー領域のゲノム比較。区間全体で遺伝子の順序は保存されている。*Pw* の遺伝子座は 7 つの *Fth* コピーを持つのにに対し、オーソログな *Xt* 領域は 4 つを含む。

(b) *Fth* クラスターに隣接する領域も強いシンテニーを示す。シンテニーブロックは *Xt* では 600 kb だが、*Pw* では 4.7 Mb に拡大している (約 8 倍)。クラスター内部では、*Xt* の 50 kb の領域が *Pw* では 770 kb に拡大しており、これは一部コピー数増加を反映した約 15 倍の拡大である。

(c) *Pw* と *Xt* のフェリチン遺伝子の最尤系統樹。祖先型 *Fth1* 系統をハイライトしている (パネル a および c の緑丸)。

この領域にはフェリチン重鎖ファミリー (*Fth*) 遺伝子クラスターに加え、近傍にある 12 個のオーソログ近傍遺伝子 (*Kcn7*, *Kif18b*, *Ntf4*, *Saxo3*, *Cgb3*, *Ruvbl2*, *Gys1*, *Ras1*, *Ppp1r15*, *Nucb1*, *Tulp2*, *Dhhdh*) が含まれていた (図 4a)。*Xt* ゲノムでは、4 つの *fth* 遺伝子が約 50 kb の区間に存在するのにに対し、*Pw* ゲノムでは 7 つの *Fth* 遺伝子が約 770 kb という拡大した領域を占めていた (図 5)。

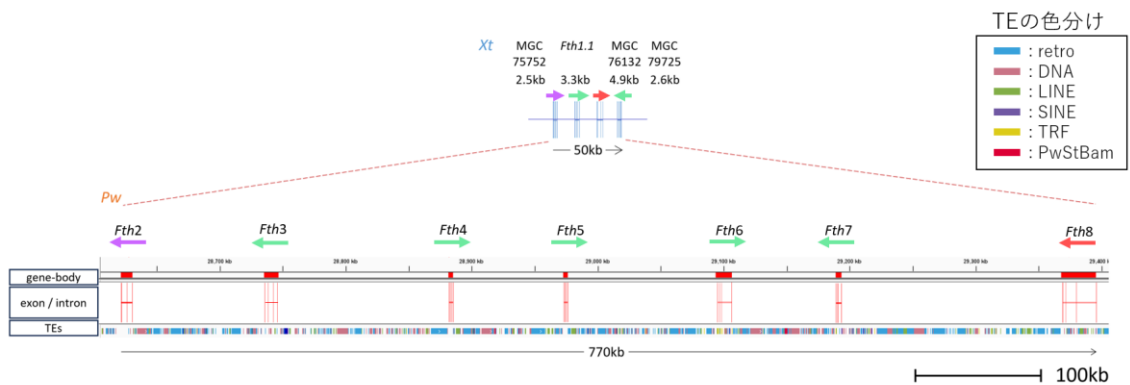


図 5. *Xt* と比較した *Pw* におけるフェリチン (*Fth*) 遺伝子クラスターの拡大と TE に富む構造 *Xt*: 上段と *Pw*: 下段間の *Fth* 遺伝子座の比較。 *Xt* では 4 つの *fth* 遺伝子が 50 kb の区間に存在するが、 *Pw* では 7 つの *Fth* 遺伝子 (*Fth2-Fth8*) が拡大した 770 kb の領域に同定される。 *Pw* 遺伝子座の下のトラックは、遺伝子領域(赤: gene body)の境界、エクソン-イントロン構造、および TE の分布(色は図 3 に準拠)を示しており、クラスター全体にわたる広範な反復配列の蓄積と、 *Pw* における領域拡大を強調している。

シンテニーブロック全体では、 *Xt* の約 0.6 Mb に対し、 *Pw* では約 4.7 Mb に拡大していた (図 4b)。すなわち、 *Pw* における *Fth* 重鎖遺伝子クラスターは、コピー数が 4 から 7 へ増加したことに伴い、領域として約 15 倍 (50 kb → 770 kb) の拡大を遂げていた。クラスターの外側では、オーソログな遺伝子を含むシンテニー領域全体が 8 倍に拡大 (0.6 Mb → 4.7 Mb) しているにもかかわらず、隣接遺伝子の数と順序は種間で概ね保存されていた。

ヤツメウナギ (*Petromyzon marinus* : *Pm*) の *fth B-like* 遺伝子を外群とした分子系統解析 (図 4c) の結果、 *Pw* の *Fth2* と *Fth8* はそれぞれ *Xt* の *MGC75752* および *MGC76132* とオーソログであることが示された。対照的に、 *Pw* の *Fth3-7* および *Xt* の *fth1*、 *MGC79725* は、祖先型 *Fth1* 遺伝子の系統特異的な重複によって生じ

たと考えられる (図 4a, c の緑丸)。 *Fth1* 由来遺伝子のコピー数は *Pw* の進化過程で約 4 倍に増加したと推定され、これが *Fth* 領域の局所的拡大に大きく寄与していることがわかる。

***Pw* の *Fth* 領域における拡大とトランスポゾンの蓄積**

Pw におけるゲノム拡大の実質的な要因は、TE の広範な挿入に帰せられる (図 6)。図 6a において、上段は 800 kb の領域にわたる *Fth2-8* 遺伝子領域の分布を示している。中段は肝臓トランスクリプトーム由来のリードマッピングを示し、各 *Fth2-8* 遺伝子に 4 つのエクソンの領域があることを示している。下段は反復配列の分布を描写しており、レトロトランスポゾン (retro; シアン)、DNA トランスポゾン (マゼンタ)、LINE および SINE (黄緑および紫)、タンデムリピート (TR; 黄)、イモリ特異的サテライト配列 (赤) が含まれる。注目すべきは、*Fth* 遺伝子間の遺伝子間領域全体にわたり、これらの TE がほぼランダムにかつ高頻度で挿入されている点である。

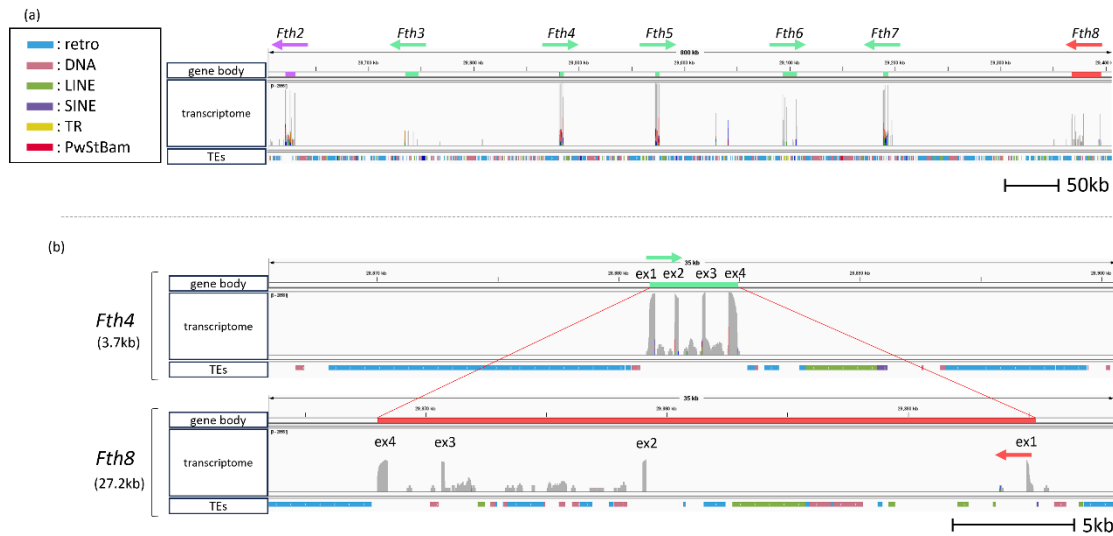


図 6. *Pw Fth2-Fth8* 遺伝子座における遺伝子領域構造、転写産物から得られたエクソン-イントロン構造、およびリピート組成

(a) *Fth2-Fth8* ゲノム領域の概要。上段は *Fth2-Fth8* のエクソン-イントロン構造を示す (図 1 と同様に色分け)。中段はエクソン境界を支持するトランスクリプトームリードのマッピングを示す。下段は反復因子の組成をクラスごとに色分けして示す：シアン、レトロトランスポゾン (retro)；黄緑、LINE；紫、SINE；マゼンタ、DNA トランスポゾン；赤、イモリ特異的サテライト (PwStBam)；黄、タンデムリピート (TR)。

(b) *Fth4* (コンパクト、3.7 kb) と *Fth8* (肥大、27.2 kb) の遺伝子領域比較。両者とも 4 エクソン構造を保持しているが、*Fth8* はイントロンへの多数の TE 挿入を含んでおり、約 10 倍の肥大化を引き起こしている。両方の遺伝子座とも隣接領域に TE を含んでいる。

Fth4 と *Fth8* 領域の拡大図である図 6 が示す通り (上下パネル、同縮尺)、*Fth4* 遺伝子領域 (約 3.7 kb) にはイントロン内に TE が存在しないのに対し、*Fth8* 遺伝子領域のイントロン内には多様なタイプの TE が挿入され、サイズが約 3.7 kb から 27.2 kb へと拡大している。しかしイントロンへの多種多様な TE の挿入によって遺伝子領域が大幅に巨大化しているにもかかわらず、トランスクリプトームデータの mRNA 配列を詳細に調べるも、フェリチンをコードする

完全なオープンリーディングフレーム (ORF) が適切に維持されていることが示された (図 7a)。さらにアミノ酸配列の比較からも、これらの挿入がフェリチンタンパク質の機能は損なわれていないことが示唆された (図 7b)。



図 7. TE により肥大化した *Fth8* 遺伝子におけるフェリチンコード ORF の保存

(a) *Pw* の FTH4 (上) および FTH8 (下) のトランスクリプトームから得られたスプライシング済み mRNA 配列を翻訳した予測アミノ酸配列。スプライシングされたエクソン配列を IGV での RNA-seq リードに基づいて抽出し、オンライン ORF/タンパク質予測ツールを用いて翻訳した。*Fth8* ゲノム領域のイントロン内に広範な TE 挿入があるにもかかわらず (図 6b)、成熟 *Fth8* 転写産物は連続したタンパク質コード ORF を保持している。

(b) 予測された FTH4 および FTH8 アミノ酸配列のペアワイズアライメント。フェリチンコード領域全体で高い配列保存性を示しており、拡大した *Fth8* 遺伝子座において CDS が保存されていることと一致する。

Fth2 から *Fth8* までの遺伝子領域長 (bp) と反復配列含有率 (%) の関係を図 8 にまとめた。*Fth2*、*Fth3*、*Fth6*、および *Fth7* のイントロン内には、主にレトロトランスポゾンと DNA トランスポゾンの両方からなる多数の TE 挿入が検出された。またこの遺伝子群において、遺伝子領域が長いほど反復配列の含有率が高

い傾向にあった (図 8b)。イントロン内に様々な TE 挿入が存在するにもかかわらず、スプライシングされた mRNA は、*Fth8* 遺伝子で見られたように、ORF が保持されていた (図 9)。同様の TE 主導の拡大は *Fth* 領域に隣接する領域でも観察されるが、フェリチン遺伝子以外の数と順序は概ね保存されていた (図 4a)。

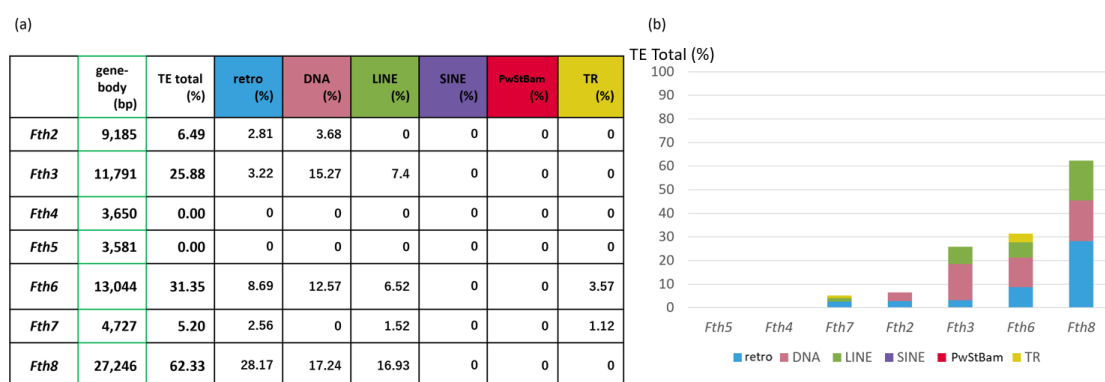


図 8. *Pw Fth* 遺伝子のリピート組成と遺伝子領域長

(a) *Pw Fth2-Fth8* の遺伝子領域長 (bp) とリピート含有率の要約表。TE total (%) はトランスポゾン (TE) としてアノテーションされた遺伝子領域の割合を示し、クラス別値 (%) は遺伝子領域内における各リピートクラスの寄与を示す。

(b) 各 *Fth* 遺伝子領域内のリピート組成 (%) を TE クラス (レトロトランスポゾン (retro)、DNA トランスポゾン、LINE、SINE、サテライト、およびタンDEMリピート (TR)) ごとに分割して示した積み上げ棒グラフ。ファミリーメンバー間での TE 蓄積の比較を容易にするため、遺伝子はプロット内で順序付けられている。

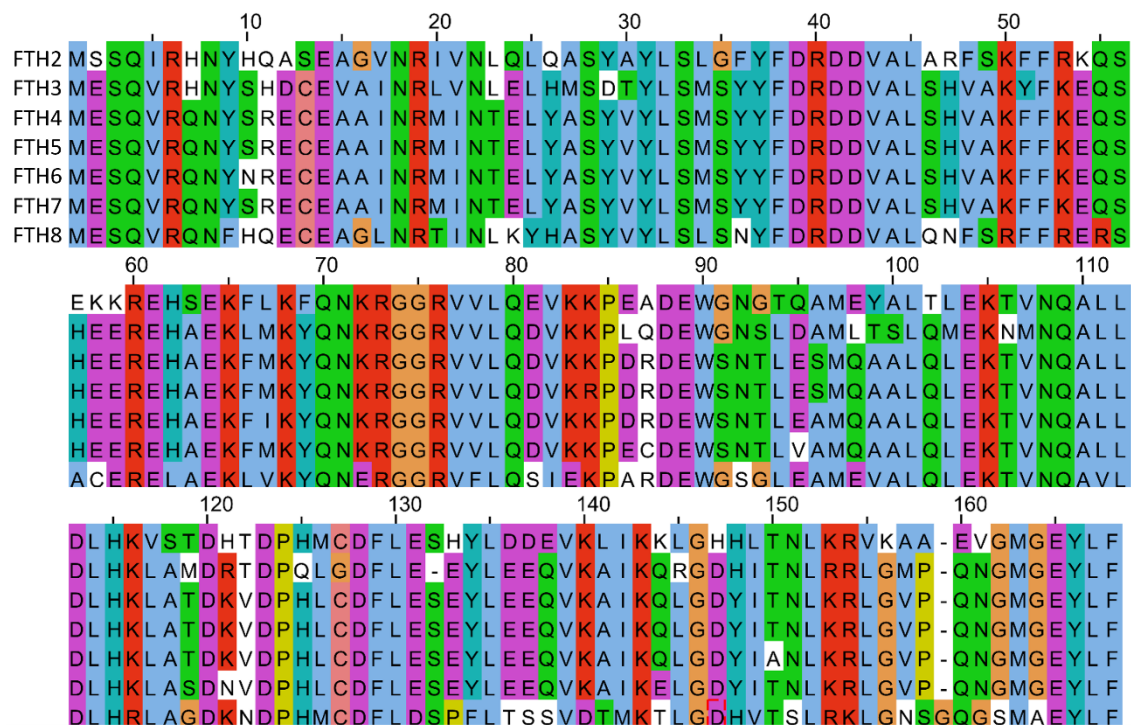


図 9. *Pw* FTH2-FTH8 予測フェリチンタンパク質のアミノ酸アライメント

転写産物から得られたコード領域から翻訳したアミノ酸配列を MAFFT を用いてアライメントし、Jalview で物理化学的性質ごとに残基を色分けして可視化した。アミノ酸配列の比較により、いくつかの遺伝子座で広範なイントロンへの TE 挿入があるにもかかわらず、7つのタンパク質すべてが高い保存性を示すことが明らかになった。これはフェリチンコード ORF が保存されていることと一致する。

CAGE/NET-CAGE 解析は TE 挿入により肥大化した遺伝子の転写抑制を示す

これまで述べてきたようにトランスクリプトーム、CAGE、および NET-CAGE 解析では、肝組織において *Fth3*, *Fth6*, *Fth8* からの転写シグナルはほとんど検出されなかった (図 10a)。これは、ORF の配列は正しく保たれていると考えられるにもかかわらず、これらの遺伝子の転写活性が著しく低下していることが示している。さらに *Fth* 遺伝子ファミリー全体で見ると、転写活性は遺伝

子サイズと逆相関を示し、TEの蓄積によって巨大化した遺伝子ほど発現レベルが低下していた(図10b)。これらの解析結果から、一部の重複遺伝子は多コピーへの増幅の後、遺伝子領域および隣接領域への進行的なTE挿入の結果として転写的に不活性化されることが示唆された。

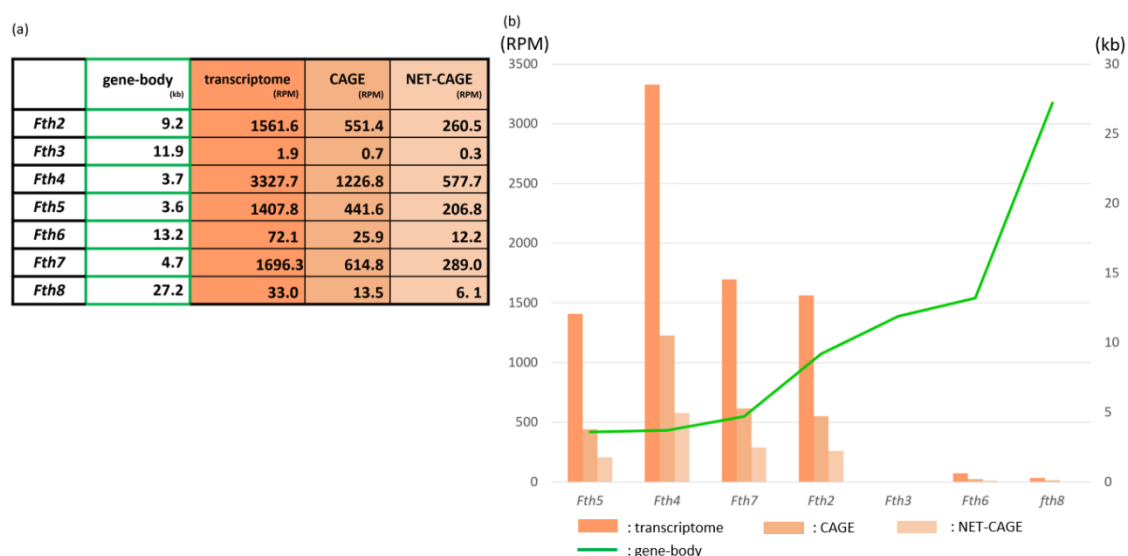


図10. フェリチン遺伝子ファミリーにおいて遺伝子領域サイズは転写量と逆相関する

(a) トランスクリプトーム、CAGE、およびNET-CAGEデータセットに基づく成体肝臓における *Pw Fth2-Fth8* の発現レベル (RPM)。各遺伝子の遺伝子領域サイズ (kb) も示す。

(b) パネル(a)のグラフ表現。遺伝子は遺伝子領域サイズの昇順(緑色の線)で並べられている。各遺伝子について、3本の棒グラフはトランスクリプトームRPM(左)、CAGERPM(中)、NET-CAGERPM(右)を示す。遺伝子領域サイズが増加するにつれて転写量が著しく低下しており、イントロンへのTE挿入が遺伝子領域を拡大させ発現を減少させることと一致している。

Alb 遺伝子座におけるTEの挿入による遺伝子サイズの肥大化と転写活性の低下

先述の傾向をさらに評価するため、通常肝組織で高発現しているアルブミン(*Alb*) 遺伝領域について解析した。*Alb* 領域では、*Pw* と *Xt* の双方で3つのアルブミンファミリー遺伝子 (*Alb1*, *Alb2*, *Alb3*) が同定された。*Xt* においてこれら

の遺伝子はそれぞれ約 16.8 kb、50.9 kb、12.3 kb の範囲に収まっている (図 11)。対照的に、*Pw* の対応遺伝子はイントロン領域への広範な TE の挿入により拡大しており、コピー数は変わらないものの、それぞれ約 62.5 kb、160.9 kb、435.7 kb と大幅に拡大していた (図 11)。

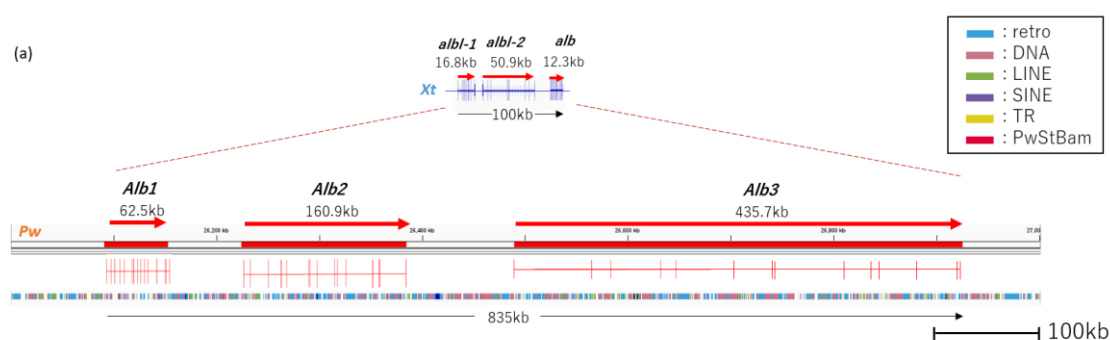


図 11. *Xt* と比較した *Pw* におけるアルブミン (*Alb*) 遺伝子クラスターの TE による遺伝子領域の巨大化

(a) *Xt* (上) と *Pw* (下) 間のアルブミン遺伝子座の模式的比較。3つのアルブミンファミリー遺伝子 (*Alb1*, *Alb2*, *Alb3*) が両種に存在し、コピー数とシンテニーは保存されている。各遺伝子の上に遺伝子領域長を示す (*Xt*: 16.8 kb, 50.9 kb, 12.3 kb; *Pw*: 62.5 kb, 160.9 kb, 435.7 kb)。 *Pw* 遺伝子座の下のトラックは、遺伝子領域の境界、エクソン-イントロン構造、および TE の分布を示しており、特に *Alb3* において著しい遺伝子領域拡大に関連した広範なイントロンへの TE 挿入を強調している。

TE 挿入により *Pw* の *Alb3* 遺伝子領域が 30 倍以上に拡大しているにもかかわらず、エクソン数は 15 個のみであり、完全な ORF が保存されている点である (図 12a)。さらに、*Alb1* にコードされる ALB1 とのアミノ酸配列比較によって、ALB3 アルブミンタンパク質の機能が概ね保存されていることが示唆された (図 12b)。

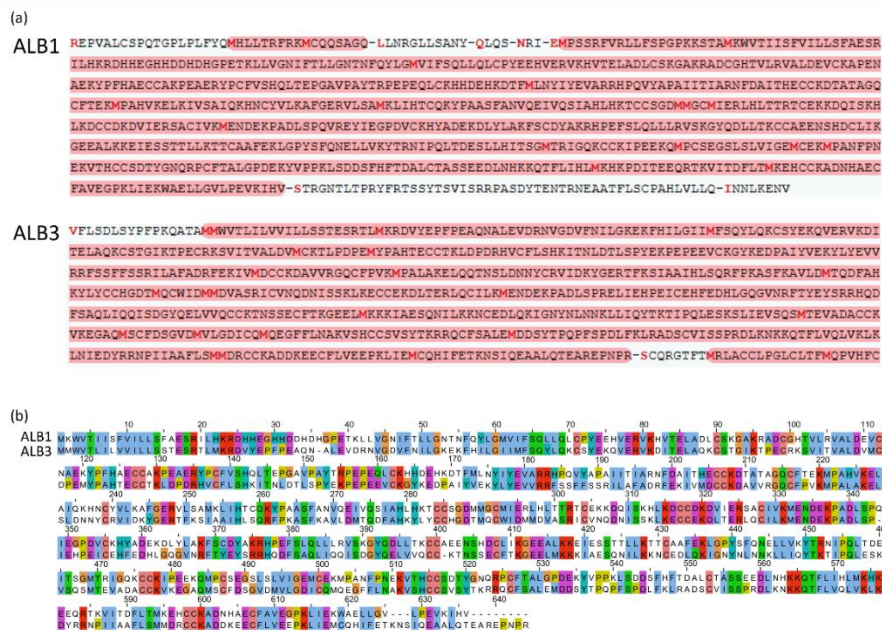


図 12. *Pw* の ALB3 における極端な遺伝子領域拡大にもかかわらず保存されたアルブミン

(a) *Pw* の ALB1 (上) および ALB3 (下) のトランスクリプトームから得られたスライシング済み mRNA 配列から翻訳された予測アミノ酸配列。スライシングされたエクソン配列を IGV での RNA-seq リードに基づいて抽出し、オンライン ORF/タンパク質予測ツールを用いて翻訳した。イントロンへの TE 挿入により ALB3 遺伝子領域が大幅に肥大化しているにもかかわらず (図 8)、成熟した *Alb3* 転写産物は連続したタンパク質コード(ORF)を保持している。

(b) 予測された ALB1 および ALB3 アミノ酸配列のペアワイズアライメント。アルブミンコード領域全体で広範な配列保存性を示しており、拡大した *Alb3* 遺伝子座においてアミノ酸コード能が保存され、タンパク質機能が維持されている可能性が高いことと一致する。

この肥大化と一致して、mRNA 量および CAGE/NET-CAGE シグナル強度は、*Fth* 遺伝子でみられたように遺伝子領域サイズに逆比例して減少していた (図 13)。

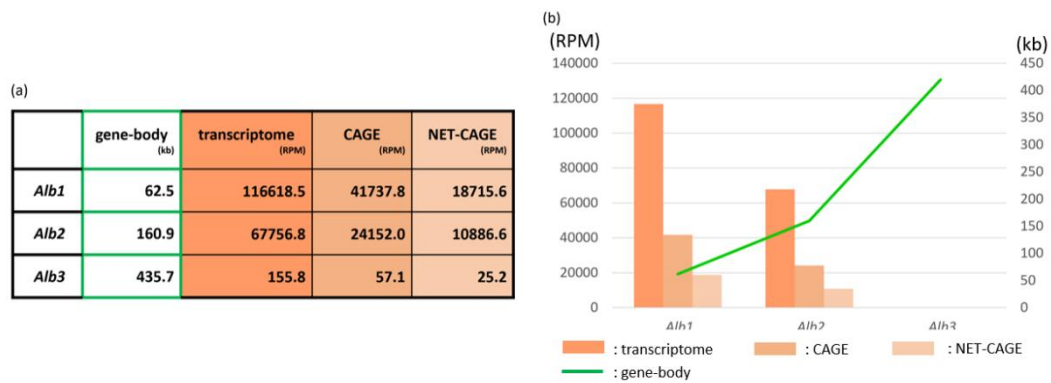


図 13. *Alb* 遺伝子ファミリーにおいて遺伝子領域サイズは転写量と逆相関する

(a) トランスクリプトーム、CAGE、および NET-CAGE データセットに基づく成体肝臓における *Pw Alb1-Alb3* の発現レベル (RPM)。

(b) パネル(a)のグラフ表現。順序とプロット形式は図7の形式に従う。より大きな遺伝子領域は転写の減少と関連している。

Vtg 遺伝子座における TE の挿入による遺伝子転写活性の低下と偽遺伝子化

この傾向をさらに調べるため、肝組織で高発現している卵黄タンパク質のビテロジェニン (*Vtg*) の遺伝子領域についても詳細に解析した。その結果、以下に述べる通りの傾向が *Vtg* 遺伝子領域で認められた。*Vtg* 遺伝子領域については *Pw* ゲノム内に 12 個の *Vtg* 遺伝子を同定した (図 14)。*Xt* の対応する 3 つの *vtg* 遺伝子は平均約 17.7 kb (14.9 kb, 18.2 kb, 20.0 kb であったのに対し、*Pw* の *Vtg1* から *Vtg11* の遺伝子領域サイズは約 59.3 kb から 80.8 kb の範囲) である。中でも *Vtg12* はより高度な TE の蓄積により約 378.9 kb へと拡大していることが判明した。

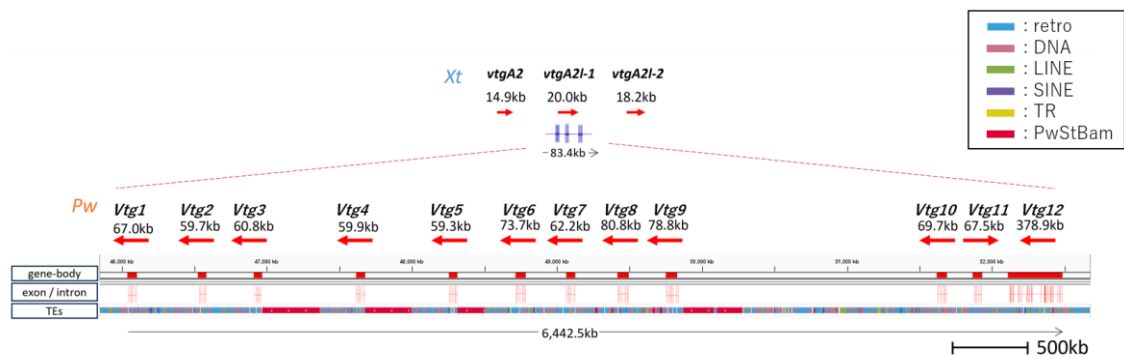


図 14. *X. tropicalis* と比較した *P. waltil* におけるピテロジェニン (*Vtg*) 遺伝子クラスターの拡大と TE に富む構造

Xt (上) と *Pw* (下) 間の *Vtg* 遺伝子座の模式的比較。*Xt* には 3 つの *Vtg* 遺伝子が存在するが、*Pw* では 12 個の *Vtg* 遺伝子 (*Vtg1-Vtg12*) が同定された。*Pw* 遺伝子座の下のトラックは、遺伝子領域の境界、エクソン-イントロン構造、および TE の分布を示しており、クラスター全体にわたる広範な反復配列の蓄積と、複数の *Pw Vtg* コピーにおける著しい遺伝子領域の肥大化を強調している。

予想どおり、378.9 kb と TE の挿入によって巨大化した *Vtg12* は高発現している 80kb 前後の *Vtg8* や *Vtg9* 遺伝子と比べて転写量が低下していた (図 15)。しかしながら、これだけ肥大化したにもかかわらず 35 個のエクソン-イントロン構造と ORF が保持されていた (図 16a)。また、遺伝子サイズが 60 kb 前後の一部の *Vtg* 遺伝子に著しい転写活性の低下が観察された (図 12)。特に *Vtg6* は、*Vtg12* ほど肥大化していないものの、ほぼ完全な転写活性の停止が見られた (図 15a)。さらに *Vtg6* の ORF を解析したところ、その ORF は既に部分的に変異が認められ、タンパク質をコードするというゲノムの機能の進行的な退化が示された (図 16b)。

これらの解析は、イントロンへの TE 挿入がまず遺伝子領域の巨大化

(inflation) と転写活性の低下を引き起こし、続いて機能喪失変異の蓄積、そして最終的な偽遺伝子化(pseudogenization)へと至るという道筋を示唆した。

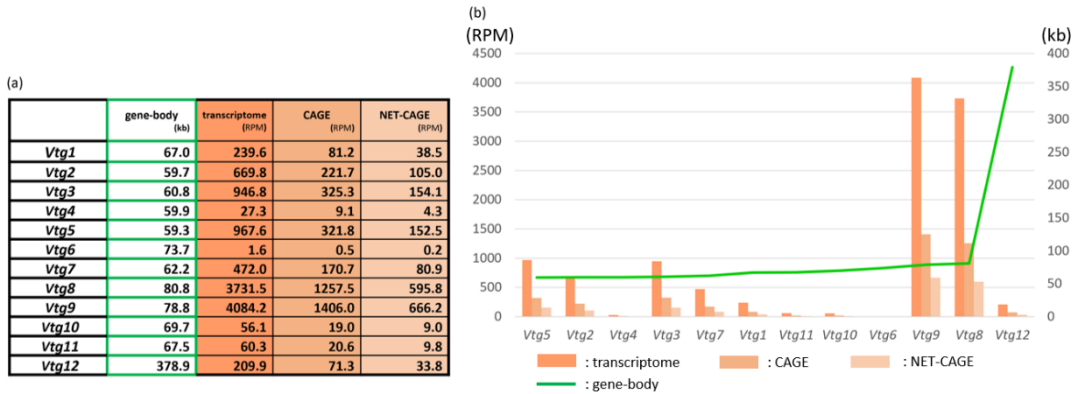


図 15. *Vtg* 遺伝子ファミリーにおける遺伝子領域サイズと転写量の関係

(a) トランスクリプトーム、CAGE、および NET-CAGE データセットに基づく成体肝臓における *Pw Vtg1-Vtg12* の発現レベル (RPM)。各遺伝子の遺伝子領域サイズ (kb) も示す。
 (b) パネル(a)のグラフ表現。遺伝子は遺伝子領域サイズの昇順 (緑色の線) で並べられている。各遺伝子について、3本の棒グラフはトランスクリプトーム RPM (左)、CAGE RPM (中)、NET-CAGE RPM (右) を示す。高発現遺伝子 (*Vtg8*, *Vtg9*) と、転写量が低下した著しく拡大した *Vtg12* 遺伝子とは対照的である。

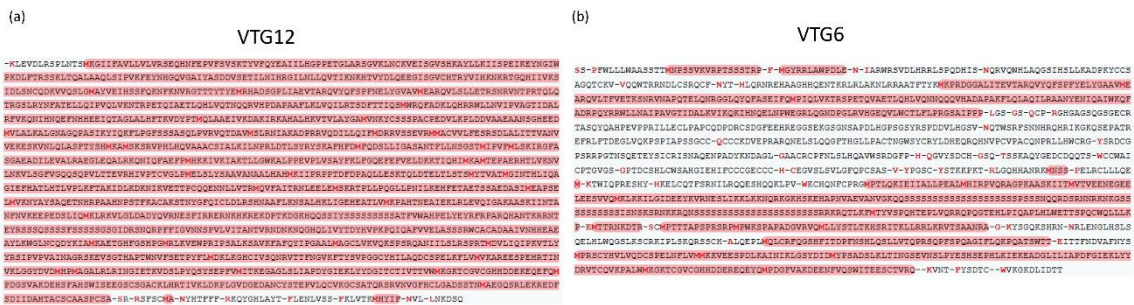


図 16. *Pw* の拡大した *Vtg* 遺伝子座間の異なるコード配列

(a) トランスクリプトーム支持スプライシング済み mRNA 配列から翻訳された *Vtg12* の予測アミノ酸配列。極端な遺伝子領域拡大にもかかわらず、連続したタンパク質コード ORF が保持されていることを示す。
 (b) トランスクリプトーム支持スプライシング済み mRNA 配列から翻訳された *Vtg6* の予測アミノ酸配列。タンパク質コード能の進行的な喪失と一致する、ORF の部分的な破壊 (例：フレームシフトや早期終止コドン) を示す。

Vtg 遺伝子クラスター領域における TE による一部 Vtg 遺伝子の破壊

Vtg9 遺伝子と Vtg10 遺伝子 の間の領域には、Vtg ファミリー遺伝子と相同な多数のエクソン様フラグメントが散在しているのが見つかった (図 17a、青枠)。これは、TE の挿入・転移の繰り返しによって、かつて完全であった遺伝子領域を分断または破壊した可能性を示唆している。Vtg ファミリー遺伝子のエクソン数は、大部分の Pw Vtg (12 個) および Xt vtg (3 個) において 35 個として保存されていた。

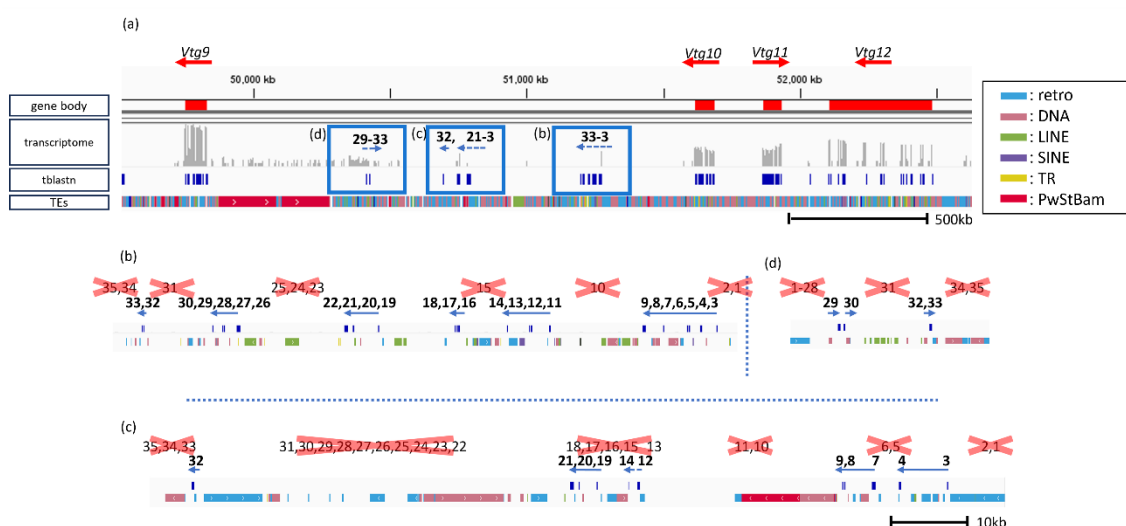


図 17. Vtg9 と Vtg10 遺伝子座間における TE に関連した Vtg エクソン構造の破壊

(a) Vtg9-Vtg12 にまたがる Vtg ゲノム領域の概要。トラック (左) は遺伝子領域区間、肝臓トランスクリプトームリードマッピング、Xt の vtgA2 由来タンパク質配列をクエリとした tblastn ヒット、および反復要素の分布を示す。リピートクラスは色分けされている (右)：レトロトランスポゾン (retro)、シアン；DNA トランスポゾン、マゼンタ；LINE、緑；SINE、紫；タンデムリピート (TR)、黄；およびイモリ特異的サテライト PwStBam、赤。青枠は Vtg 遺伝子と相同な分散したエクソン様断片を含む 3 つの領域を示す。

(b-d) (a) の枠で囲まれた領域の拡大図。Xt vtgA2 の 35 エクソン構造にアライメントされたエクソン様断片を示す。保存されたエクソン番号を黒で示す。各破壊領域から欠失しているエクソンには赤い「X」マークを重ねて示している。

機能喪失に至る変異が認められた遺伝子本体の具体例として、図 17b は合計 10 個のエクソン（エクソン 1-2, 10, 15, 23-25, 31, 34-35）が失われていた（図 17b）。図 17c では 24 個のエクソンが欠失していた（エクソン 1-2, 5-6, 10-11, 13, 15-18, 22-31, 34-35）（図 17c）。図 17d ではさらに極端なパターンが観察され、エクソン 1 から 28 が完全に消失し、わずか 4 つのエクソン（29, 30, 32, 33）のみが残存していた（図 17d）。マッピング解析により、これらの欠失領域は LINE、レトロトランスポゾン、DNA トランスポゾンを含む大型 TE によって広範囲に侵入されていることが明らかになった（図 17b-d）。図 17d の破壊された遺伝子領域では、エクソン 31 に相当する領域に複数の LINE 挿入が見られ（図 17d）、欠失したエクソン 1-28 にまたがる上流領域は、イモリ特異的サテライト配列（PwStBam、図 17a の赤色）の巨大なブロックに置き換わっていた。これらの散在するエクソン様配列は、遺伝子をコードするという機能を失った、「死んだ」遺伝子の残骸である可能性が高い。

これらの結果は、TE が *Pw* ゲノム全体で活発に転移し、遺伝子間領域およびイントロン領域へ広く挿入していることを示している。このような広範かつ大部分が確率論的な挿入は、遺伝子領域を巨大化させることに止まらず、最終的に影響を受けた遺伝子座の転写停止または機能破壊を惹起したと推察された。

Pw ゲノムにおける最近転移したレトロトランスポゾンの検出

日本の *Pw* 近交系（以下 JP *Pw* ゲノム）と Simon らのグループが報告した参照ゲノム（EU *Pw* ゲノム）との間で *Fth* 領域を比較した際、Harr プロットと VISTA を用いて（図 18a, b）、JP *Pw* ゲノムにのみ存在するレトロトランスポゾン挿入を *Fth5* の 3' 下流領域に同定した（図 18b, *1, 赤色ハイライト）。挿入されたエレメントは約 5.9 kb の長さを持ち、LTR を含む *Gypsy* ファミリーのレトロトランスポゾンに属していた（図 18c, *1）。

この主要な挿入に加え、JP *Pw* ゲノムの同領域内で他に 4 つの挿入イベントが検出された（図 18b, c）。2 つ目は 5' LTR の U3 領域に由来する短い 193 bp の断片からなり（図 18c, *2）、レトロトランスポゾンが組み込まれた後、その部位から二次的な転移を起こした可能性を示唆している。3 つ目は、両側の TIR（末端逆位配列）を含む残存した 594 bp の Harbinger 型 DNA 断片からなり、DNA トランスポゾンが組み込まれた後、そこから二次的な転移を起こしたことを示唆している（図 18c, *3）。4 つ目は、*Fth6* イントロンの高密度反復配列領域内に挿入された非常に小さな 42 bp の断片である（図 18c, *4）。ただし、イントロンへの挿入がわずか 42 bp であったため、JP サンプルのいずれにおいても肝臓での *Fth6* 遺伝子の発現にはほとんど影響が見られなかった。

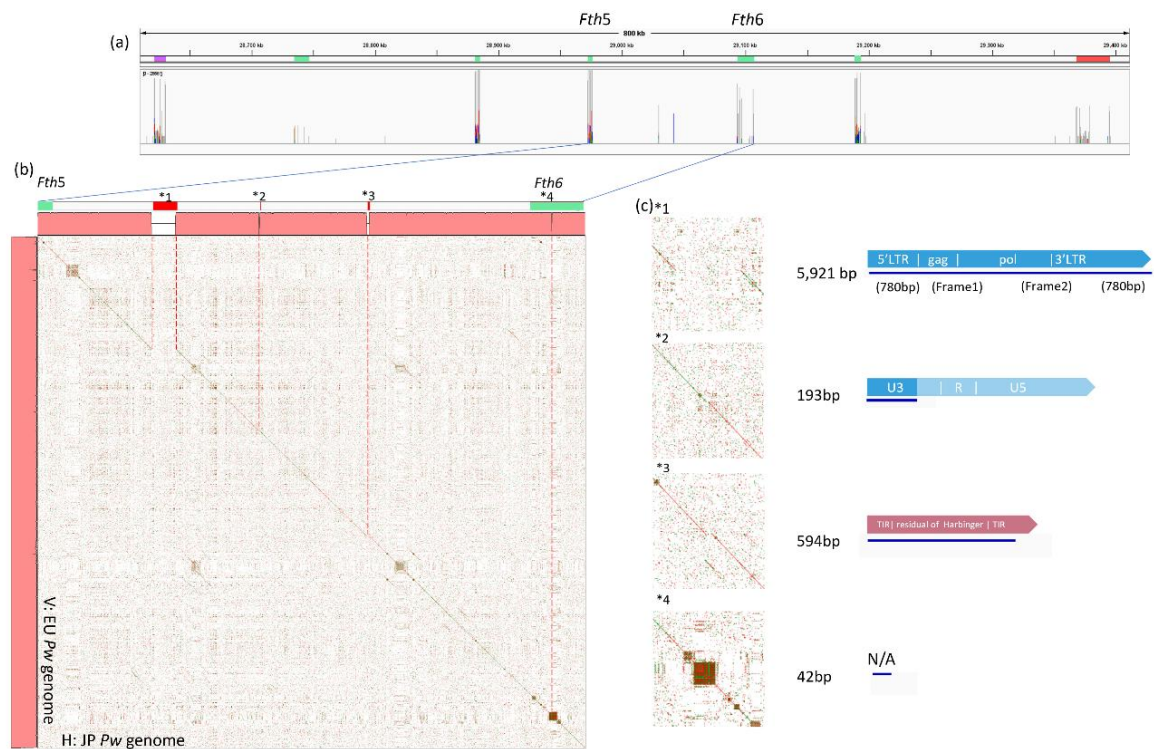


図 18. EU-JP 系統間のゲノム比較により明らかになった *Fth* 領域における新規挿入トランスポゾン

- (a) Harr プロットと VISTA によって解析された *Fth* 領域 (*Fth5* と *Fth6* の間、約 135 kb) の概要。
 (b) EU 系統と JP 系統の比較アライメントにより、EU ゲノムには存在しない JP 特異的な TE 挿入(*1-*4)を同定。Harr プロットと VISTA の結果を組み合わせている。JP および EU ゲノムはそれぞれ水平軸および垂直軸に配置されている。
 (c) 挿入部位における Harr プロットの拡大図。右パネルは挿入された TE の構造とサイズを示す。

これらの観察は、*Gypsy* 様レトロトランスポゾン及び、恐らく他のいくつかの TE が、EU と JP が解析したそれぞれのイモリの祖先が分岐されたとされる 100 万年の間に転移イベントを起こしたことを示している。このような現在進行形の活動は、レトロトランスポゾンが *Pw* ゲノムを動的に拡大させ続けているという見解をさらに支持するものである。

***Pw* 精巣で高発現するユニークな構造を持つレトロトランスポゾンの発見**

卵母細胞がランプブラシ染色体に代表される開放的クロマチンをもつことから、当初 *Jp* 系統で検出された *Gypsy* 様レトロトランスポゾンが卵巣で転写活性を持ち、それが最近の転移イベントを引き起こした可能性があるという仮説を立てた。しかし、トランスクリプトーム解析では、3ヶ月齢および7ヶ月齢の卵巣、未受精卵、精巣のいずれにおいても *Gypsy* 様エレメントの発現は検出されなかった。そこで、卵巣と精巣で発現している他の逆転写酵素 (RT) コード配列を探索したところ、3ヶ月齢の精巣で強く発現し、卵巣では弱い発現にとどまる RT 遺伝子が同定された (図 19a, b)。その転写産物は約 5.5 kb の長さで、DIRS1 型逆転写酵素と RNase H 様ドメインの両方をコードしている (図 20a)。想定外なこととして、別の読み枠 (alternative reading frame) において、ファージインテグラーゼ様およびチロシンリコンビナーゼ (YR) 様タンパク質がコードされていた (図 20b)。

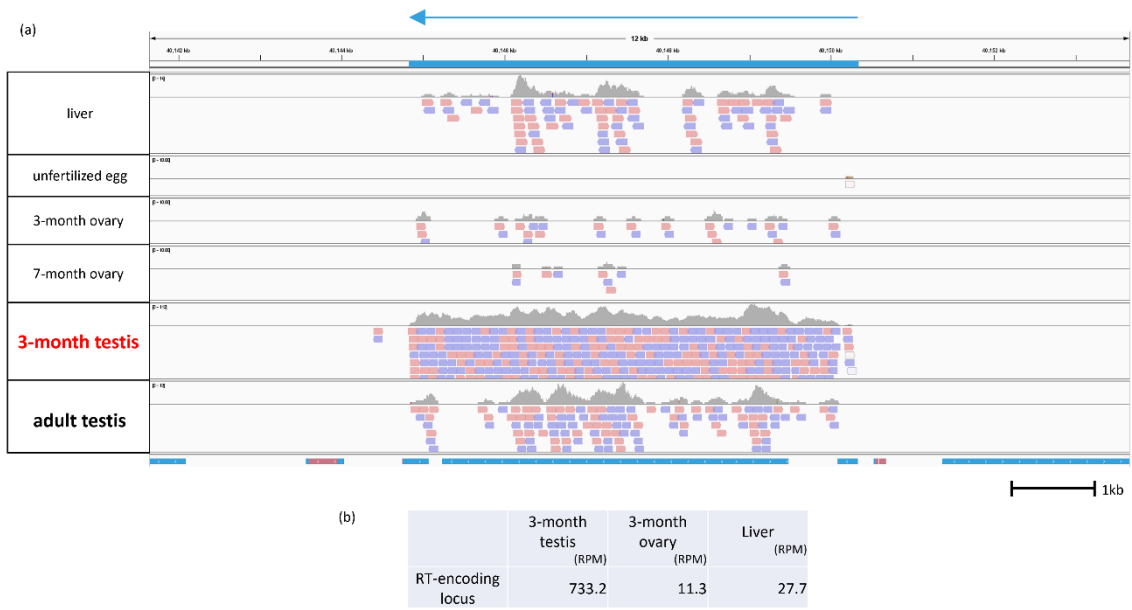


図 19. Pw における RT コード遺伝子の精巣特異的な発現

(a) 複数の組織/段階（肝臓、未受精卵、3 ヶ月齢卵巣、7 ヶ月齢卵巣、3 ヶ月齢精巣、成体精巣）における RT コード遺伝子座の RNA-seq リードカバレッジとアライメントのゲノムブラウザ表示。精巣（特に 3 ヶ月齢精巣）において特異的な転写が観察される一方、卵巣や未受精卵ではシグナルがほとんどまたは全く検出されない。

(b) 3 ヶ月齢精巣、3 ヶ月齢卵巣、および肝臓における RT コード遺伝子座の発現レベル (RPM) の要約。この遺伝子座は後に *Boul* と命名した。

関連する遺伝子は Xt において DNA トランスポゾンと RNA トランスポゾンの両方を動員しうる潜在的因子として報告されているが、この Xt のホモログは主に卵巣で発現し、単一の XerC 様リコンビナーゼドメインのみを含んでいる (Gazolla et al., 2022)。対照的に、 Pw の遺伝子座はファージインテグラーゼ様ドメインとリコンビナーゼ様ドメインの両方をコードし、精巣で顕著に高発現するという特徴を示している (図 19, 20)。

第三章 考察

有尾両生類におけるゲノム巨大化の動的プロセス

アホロートルのゲノム解析により、TE が有尾両生類のゲノム巨大化の主要な推進力であることが明らかにされた (Nowoshilow et al., 2018; Schloissnig et al., 2021)。この見解は、最近の *Pw* のゲノム解析によってさらに裏付けられた (Brown et al., 2025; Kimura et al., 2025)。本研究において我々の解析は、重複遺伝子のイントロンへ TE が段階的に蓄積し (図 21, ステップ①)、遺伝子領域の異常な長さへの「巨大化 (inflation)」を駆動していることを示した (ステップ②)。TE 駆動の肥大化が起こっても、pre-mRNA に転写された TE 配列がスプライシング機構によって除去されるため、遺伝子は機能的タンパク質を産生する能力を保持していると推測される。しかし、さらなる TE 挿入や欠失によって転写活性が失われると、選択圧が働かなくなり、変異の蓄積を許すことになる。極端な場合、TE による執拗な攻撃の下で、遺伝子領域が機能的な限界点、すなわち明白な「バースト (破綻)」に近づく兆候が観察された (ステップ③)。さらに、イベリアトゲイモリの 2 つの HiFi シークエンシングデータセットの比較は、遺伝子領域の肥大化が単なる歴史的遺物ではなく、有尾両生類ゲノムを能動的に彫刻し続けている現在進行形の動的プロセスであることを示唆している。

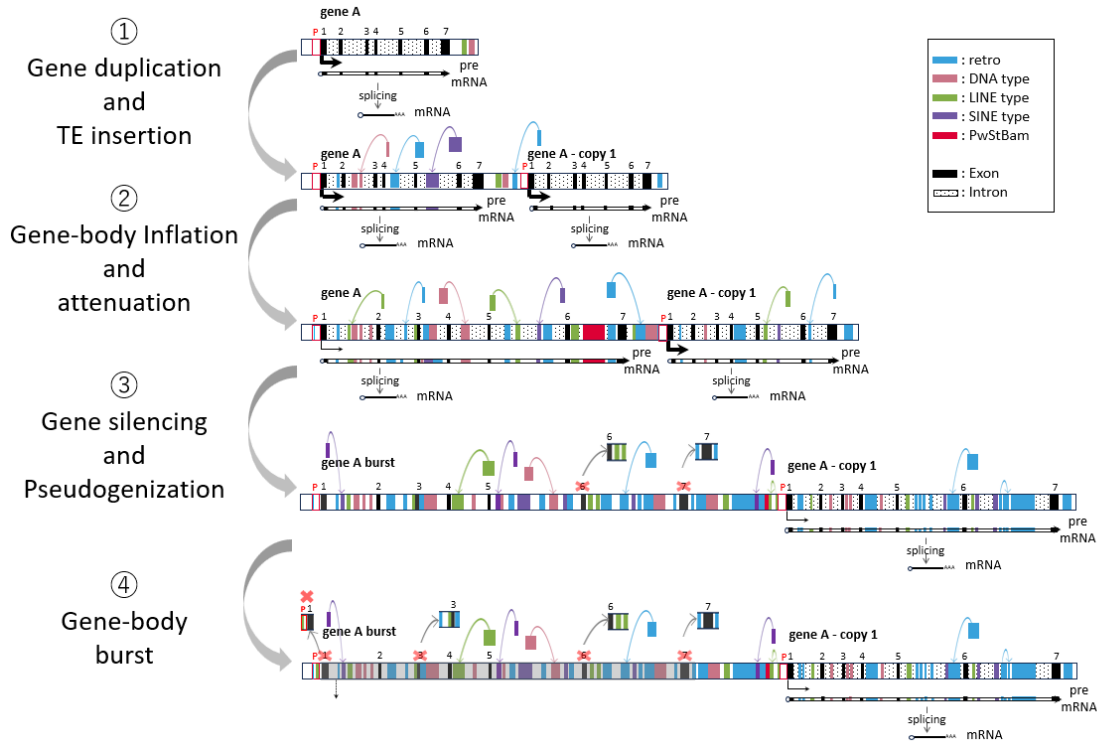


図 21. イモリゲノム拡大中の重複遺伝子における TE 主導の遺伝子領域肥大化と崩壊（「バースト」）の提唱モデル

比較解析から推測される段階的な軌跡の模式図。

ステップ 1：遺伝子重複の後、各コピーのイントロン内に TE が蓄積し、局所的なりピート密度が増加するが、エクソンの順序は保持される。

ステップ 2：継続的なイントロン性 TE 挿入が遺伝子領域を極端な長さへと進行的に肥大化（inflate）させる。ランダムな TE が転写因子結合領域などの転写を司る領域に挿入され不安定になり、転写能が低下する(attenuation)。また、TE 由来配列は pre-mRNA に取り込まれるが、スプライシングによって除去されるため、エクソン-イントロン構造が機能的である限り、ほぼ完全な成熟 mRNA の産生が可能である。

ステップ 3：持続的な TE 侵入と転写量低下後の選択圧の低下のもとで、エクソン-イントロン構造がより不安定化し、転写の不活化(silencing)と偽遺伝子化(pseudogenization)が進行する。

ステップ 4：遺伝子領域の分断/浸食（「バースト」）が起き、一方のコピーにおける遺伝子の喪失に至る。一方、他のコピーは転写を維持する場合がある。色はリピートクラス（レトロトランスポゾン (retro)、DNA、LINE、SINE、およびイモリ特異的サテライト PwStBam) を示し、黒/斜線のボックスはそれぞれエクソン/イントロンを示す。

TE 侵入に対する構造的緩衝材としてのイントロン

本研究は、50 年近く分子生物学者を魅了してきた根本的な問い、「なぜ膨大なエネルギーを消費するイントロン-エクソン構造が存在するのか？」(Berget et al., 1977; Chow et al., 1977; van den Berg et al., 1978) に私を立ち返らせた。イモリゲノムの解析は、イントロンがトランスポゾンの侵入を吸収する「構造的緩衝材 (structural buffers)」として機能し、それによってコーディングエクソンを壊滅的な破壊から守っている可能性を示唆している (図 22)。

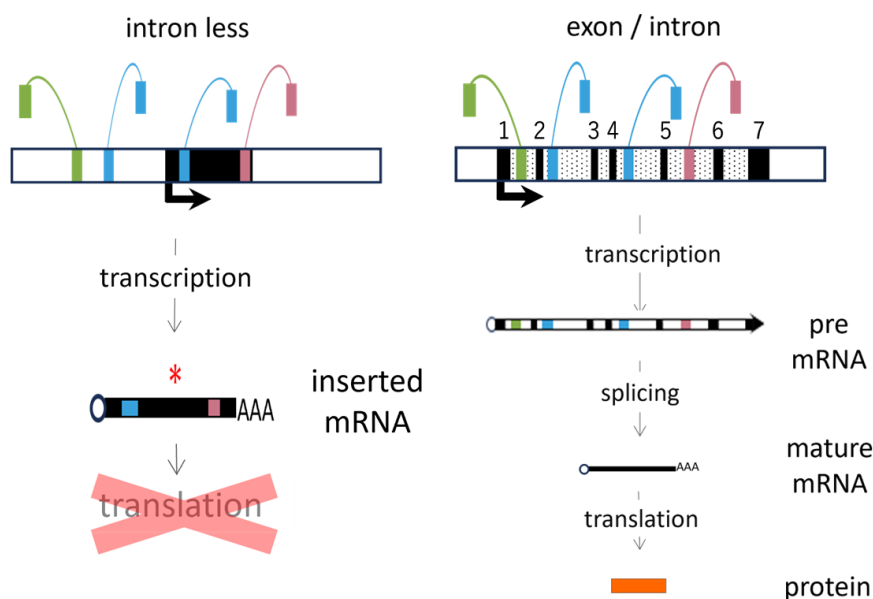


図 22. TE 挿入の破壊的影響を緩和する構造的緩衝材の模式図

イントロンTE挿入（色の付いたブロック）を受けた際のイントロンレス遺伝子（左）とエクソン-イントロン遺伝子（右）の模式的比較。イントロンレス遺伝子では、TE の組み込みがコード配列を直接破壊し、挿入を含む転写産物が生じるため、機能的タンパク質を生じる可能性は低い。対照的に、エクソン-イントロン遺伝子のイントロンに TE が挿入された場合、挿入は pre-mRNA に転写されるがスプライシングによって除去可能であり、翻訳能を維持した完全なコード領域を持つ成熟 mRNA が産生される。このモデルは、イントロンがTEの侵入を吸収しつつ、コードエクソンを壊滅的な破壊から守る仕組みを図示している。

具体的には、比較的短いコード領域のために時に 100 kb を超える長大な pre-mRNA を転写し、その後スプライシングによってイントロン由来 RNA 配列を除去することで、細胞は不可避な TE の流入を受け入れつつコーディング配列を保護するためのメカニズムとして、エクソン-イントロン構造とスプライシングシステムを獲得したのかもしれない。この防御機能を解明することで、イントロンは「ゲノムのゴミ」としてではなく、「沈黙の守護者 (silent guardians)」——真核生物ゲノムが数百万年にわたる執拗なトランスポゾンの圧力に耐えることを可能にした構造的革新——として再定義されるかもしれない。

なぜイモリでゲノム巨大化が起きたのか？

多くの生物において、Piwi/Ago システムは生殖系列内での TE 動員を防ぐ監視機構として機能しており、その結果、極端なゲノム巨大化は一般に抑制されている (Aravin et al., 2006; Siomi & Kuramochi-Miyagawa, 2009)。Pw には 3 つの *Piwi* コピー (*PiwiL1/2/4*) と 4 つの *Ago* コピー (*Ago1-4*) が存在し、すべて卵巣と精巣の両方で発現している (図 20)。したがって、他の生物と同様に、イモリの生殖系列においても Piwi/Ago 介在性の監視システムが TE 活性を効果的に抑制していると考えられる。実際、JP 系統で最近転移した Gypsy 型レトロトランスポゾンが卵巣と精巣のいずれでも発現を示さないという事実は、Piwi/Ago システムがイモリ生殖系列で機能している可能性を支持している。

(a)

	gene-body (kb)	3-month testis (RPM)	3-month ovary (RPM)	liver (RPM)
<i>Piwi1</i>	1,360.7	738.3	350.7	0
<i>Piwi2</i>	1,276.0	40.1	61.4	0
<i>Piwi4</i>	417.4	56.9	5.5	0

(b)

	gene-body (kb)	3-month testis (RPM)	3-month ovary (RPM)	liver (RPM)
<i>Ago1</i>	307.8	8.1	7.0	1.5
<i>Ago2</i>	457.3	11.5	7.4	5.2
<i>Ago3</i>	864.9	13.3	10.3	3.1
<i>Ago4</i>	452.3	14.1	35.0	2.9

図 23. *Pw* 生殖系列における *Piwi* および *Argonaute* ファミリー遺伝子の発現

(a) 3ヶ月齢精巣、3ヶ月齢卵巣、および肝臓における *Piwi* パラログ (*Piwi1*, *Piwi2*, *Piwi4*) の遺伝子領域サイズ (kb) と発現レベル (RPM)。

(b) 3ヶ月齢精巣、3ヶ月齢卵巣、および肝臓における *Ago* 遺伝子 (*Ago1-Ago4*) の遺伝子領域サイズ (kb) と発現レベル (RPM)。

さらに、本研究で同定された *Boul* 遺伝子は卵巣ではほとんど発現していないことを確認した (図 19)。しかし、なぜ *Boul* が精巣でのみ転写抑制や分解を特異的に回避しているのかを理解することが、イモリにおけるゲノム巨大化の謎を解く鍵を握っているかもしれない。

転移因子としての *Boul* の示唆に富む性質

ユニークな構造を持ち精巣に偏った発現を示す遺伝子を *Boul* と命名した。その特異な構造と生殖細胞での強い発現から、*Boul* はイモリにおいて DNA および RNA トランスポゾンが活発に動員され続けるゲノム環境に寄与している可

能性がある。その分子的機能の解明は今後の課題であるが、*Boul* が生殖系列における転移ダイナミクスに影響を与え、それによって有尾両生類のゲノム拡大のプロセスに関与している可能性が示唆された。

Boul 遺伝子は *gag* 遺伝子を欠いているにもかかわらず、DIRS 群レトロトランスポゾンに分類されてきた。DIRS 群のメンバーは、ゲノムへの組み込みにおいて、LTR 型レトロトランスポゾンや LINEs が用いる DDE 型インテグラーゼではなく、チロシンリコンビナーゼ (YR) を使用することを特徴とする (Poulter & Goodwin, 2005)。さらに、*Boul* レトロトランスポゾンは両末端に 180 bp の分断型ダイレクトリピート (SDR) 配列を含んでおり、これは *Boul* が DIRS 様エレメントとしてではなく、DIRS 群内の *Ngaro* 様レトロトランスポゾンとして分類されるべきであることを示している。

Ngaro 様レトロトランスポゾンについて、Poulter と Butler はその総説の中で「YR ORF の開始点は-1 の読み枠にある RT/RH ORF の終端とわずかに重複している。YR ORF の翻訳は、この重複領域におけるプログラムされたリボソームフレームシフトによって達成される可能性が高い」と述べている (Poulter & Butler, 2015)。同総説ではさらに、「上記のレトロトランスポゾンに加え、チロシンリコンビナーゼをコードする真核生物の DNA トランスポゾンも多数報告されている」と記されており、配列特異性は存在する可能性があるものの、YR

を含むエレメントはレトロトランスポゾンとしても DNA トランスポゾンとしても宿主ゲノムに組み込まれる能力を持っている可能性が示唆されている。

Pw ゲノムには多数の *Boul* 様レトロトランスポゾンが存在しており、その多くは精巣で優先的な発現を示す。中でも、我々が *Boullike1* と命名したエレメントは、精巣特異的な発現を示している (図 24)。

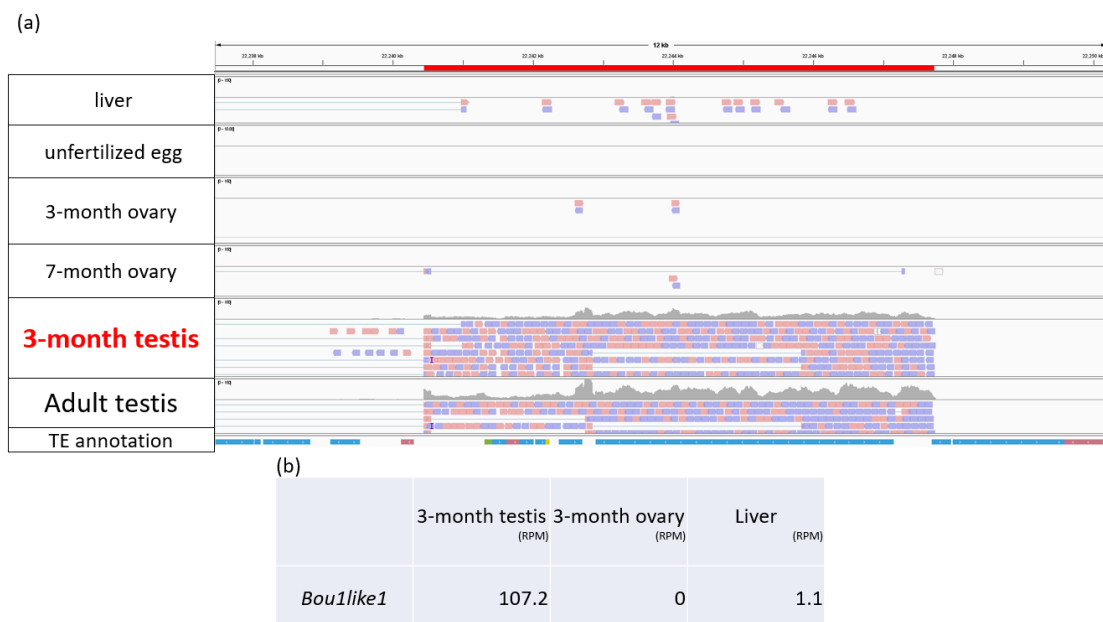


図 24. *Pw* における *Boullike1* の精巣での特異的発現

(a) 複数の組織/段階 (肝臓、未受精卵、3ヶ月齢卵巣、7ヶ月齢卵巣、3ヶ月齢精巣、成体精巣) における *Boullike1* の RNA-seq リードカバレッジとアライメントのゲノムブラウザ表示。精巣 (特に 3ヶ月齢精巣) において堅牢な転写が観察される一方、卵巣や未受精卵ではシグナルがほとんどまたは全く検出されない。

(b) 3ヶ月齢精巣、3ヶ月齢卵巣、および肝臓における *Boullike1* の発現レベル (RPM) の要約。

する抗体を作製することで、それらの時空間的な発現の詳細な解析が可能となり、TEの動員における役割について重要な知見が得られる可能性がある。さらに、*Boul*を特定のTEと共に酵母や*Xt*に導入する機能アッセイを行うことは、イントロンが実際にTEによるゲノム不安定化からコード配列を保護しているかどうかを実験的に検証するための新たな道を切り拓くと期待される。

遺伝子の **subfunctionalization** は起きているか？

遺伝子重複により複数コピーが生じた後、重複遺伝子が新たな機能を獲得し得ることはよく知られている。古典的な例として脊椎動物の水晶体におけるクリスタリン遺伝子が挙げられる。既存の代謝酵素やストレス応答遺伝子が遺伝子重複を経て、レンズという高度に特殊化した組織で水晶体タンパク質としての新規機能を獲得した。これは新機能化 (**neofunctionalization**) として知られるプロセスである (Wistow & Piatigorsky, 1988)。新たなシスエレメント獲得後の転写調節遺伝子の重複とそれに続く **subfunctionalization** も報告されている (Ochi et al., 2017)。

最近では、ハチやアリにおいてピテロジェニン遺伝子の **subfunctionalization** と **neofunctionalization** が報告されており、祖先的な発現パターンや機能とは異なる組織やカーストで発現して機能していることが報告されている (Corona et al., 2013; Salmela et al., 2022; Yaguchi et al., 2023)。

本研究では、多コピー拡大後も完全な ORF を保持しているものの、肝臓での発現が著しく低下しているイモリのビテロジェニン遺伝子を同定した。この観察から、これら重複 *Vtg* 遺伝子においても subfunctionalization が起きている可能性を検証するため、複数の発生段階と組織を網羅するトランスクリプトームデータセットを用いて *Vtg* 遺伝子の発現を調査した (Matsunami et al., 2019)。しかし、調査した利用可能なデータセットの中では、特定の組織や発生段階での明確な発現証拠は検出されなかった。現時点での我々の結果は、イモリの *Vtg* 遺伝子コピーは subfunctionalization を経ているのではなく、機能分担よりはむしろ遺伝子領域のバースト（破綻）と一致する方向へ進化していることを示唆している。

イントロンへの TE のランダムな蓄積

特定の TE が優先的にイントロンへ挿入されるかどうかを調べるため、各 *Fth* 遺伝子のイントロンに挿入された TE の種類と量を定量した。結果は色分けされた表とグラフにまとめた（図 8）。表では遺伝子は番号順に並んでいるが、グラフでは遺伝子領域サイズ順に並んでいる。これらの解析は、特定のクラスの TE がイントロン内に特異的に濃縮されているわけではないことを示している。むしろ TE 組成はゲノム全体の TE ランドスケープを反映している。これは、イントロンに挿入された TE と遺伝子間領域にある TE の間に根本的な違いはな

く、TEはイントロンと隣接領域を区別することなく、大部分ランダムな様式で転移する傾向があることを示唆している先行研究(Kimura et al., 2025)と一致する。

ゲノム巨大化と再生能力の関係

ゲノム巨大化が有尾両生類の驚異的な再生能力に寄与しているかという長年の問いに関しては、正直なところ依然として未解明であると言わざるを得ない。逆に今回の解析は、有尾両生類のゲノム巨大化は再生のための適応としてではなく、主に持続的なTE活性の副産物として生じたとする、より単純な解釈の方を支持する傾向となった。しかし、これは巨大ゲノムと再生能力との潜在的なリンクを否定するものではない。したがって、この問いを最終的に解明するためには、イモリゲノムの更なる詳細な解析が不可欠であると考えている。

実験材料と手法

イベリアトゲイモリの入手と維持

成体のイベリアトゲイモリ (*P.waltl*) は、文部科学省ナショナルバイオリソースプロジェクト (NBRP) を通じ、広島大学両生類研究センターより提供を受けた。動物はサンプリング時まで基礎生物学研究所にて飼育した。サンプル採取のため、動物は終濃度 0.1% の MS-222 (Sigma, USA) で麻酔した。すべての動物実験手順は、自然科学研究機構および広島大学の動物実験委員会の承認を得て行った。

NET-CAGE (新生 RNA) および CAGE (全 RNA) のための RNA 抽出

RNA 抽出は Hirabayashi ら (2019) の方法に従い、若干の修正を加えて実施した。*Pw* の肝組織 (150 mg) を BioMasher V チューブ (FUJIFILM Wako Chemicals) に入れ、1 ml の氷冷 Buffer A (25 μ M α -アマンニチン、1x complete Protease Inhibitor Cocktail、20 U SUPERase IN RNase Inhibitor、1 mM DTT、0.1% Triton X-100 を添加した Nuclei PURE Lysis Buffer) 中でホモジェナイズした。組織が完全に破碎された後、さらに 1 ml の Buffer A を加え、ホモジェナイズを繰り返した。

全 RNA 抽出 全 RNA 調製のため、得られたライセート 200 μ l を 700 μ l の TRI Reagent (Zymo Research) と混合し、メーカーの指示に従って Direct-zol RNA

MicroPrep キットを用いて RNA を単離した。

核分画の単離 残りのライセートを 3.6 ml の Buffer B (1 mM DTT を添加した 2 M ショ糖クッション溶液) と混合し、2 ml の氷冷 Buffer B の上に重層した。サンプルを 4°C、10,000 rpm で 60 分間遠心分離した。核ペレットを 500 μ l の Buffer C (1% TERGITOL NP-40、20 mM HEPES pH 7.5、300 mM NaCl、2 M 尿素、0.2 mM EDTA、1 mM DTT、25 μ M α -アマニチン、1x プロテアーゼ阻害剤カクテル、20 U SUPERase IN) で 1 回洗浄した。洗浄したペレットを 200 μ l の Buffer C に再懸濁し、核構造を破壊しないよう注意深くピペッティングして完全に分散させながら氷上で 10 分間インキュベートした。懸濁液を 4°C、3,000 \times g で 2 分間遠心し、ペレットを 100 μ l の Buffer C で再度洗浄した。

新生 RNA の精製 ペレットを 50 μ l の DNase I 溶液 (NIPPON GENE) で 37°C、30 分間処理した。その後 1 ml の TRI Reagent を加え、十分に混合した。ライセート量が単一の処理チューブの容量を超える場合、溶液を 2 本のチューブに等分して以後の抽出を行った。両方のチューブから Direct-zol RNA MicroPrep キットを用い、オンカラム DNase I 消化を行って新生 RNA を精製した。最初のカラムからの RNA を 30 μ l の RNase フリー水で溶出した。最終的な RNA 収量を最大化するため、この溶出液を 2 本目のカラムの溶出バッファーとして使用し、合わせた溶出液を最終的な新生 RNA サンプルとして回収した。RNA 濃度およ

び完全性は NanoPhotometer N60 (Implen) および Agilent 2100 BioAnalyzer を用いて測定した。

CAGE および NET-CAGE ライブラリ調製とシーケンシング

新生 RNA および全 RNA サンプルは DNAFORM (神奈川、日本) に提出し、CAGE ライブラリ調製およびシーケンシングを行った。RNA の完全性は BioAnalyzer (Agilent 2100) で評価した。ランダムプライマーを用いて cDNA を合成した。5'キャップ構造のリボースジオールを酸化・ビオチン化し、ビオチン化 RNA/cDNA ハイブリッドをストレプトアビジンビーズを用いて捕捉した (キャップトラッピング)。RNase ONE/H 消化および cDNA 両端へのアダプターライゲーション後、二本鎖 CAGE ライブラリを構築した。ライブラリは Illumina NextSeq 500 プラットフォームにて 75 bp シングルエンドモードでシーケンスした。

ゲノム、遺伝子アノテーション、TE アノテーション

Kimura ら (2025) によるゲノム、遺伝子アノテーション、および転移因子 (TE) データセットを使用した。同研究で報告された PwStBam 配列をクエリとして、*Pw* ゲノムに対し blastn (v2.16.0+) 検索を行った。BLAST 結果を GTF 形式に再フォーマットし、欠損フィールドはドット (".") で置換した。キュレーションされ

た配列はイモリ特異的サテライトリピートとしてゲノムアノテーションに統合した。

ソフトウェアベースの解析

リードのアライメント、転写産物の定量、系統推定、リピート解析、ドットプロット解析、およびゲノム可視化は、HISAT2 (v2.2.1), SAMtools, Salmon (v1.4.0), MEGA11 (Tamura et al., 2021), GENETYX 11 (HarrPlot3.1.0), IGV (v2.16.0), BLAST v2.16.0+, MAFFT, Jalview を含む標準的なバイオインフォマティクスツールを用いて実施した。特に記載がない限り、ソフトウェアはデフォルトパラメータで実行した。非デフォルト設定については該当する各セクションに記載した。

遺伝子発現レベルの正規化には、Reads Per Million (RPM) を算出した。CAGE および NET-CAGE データの場合、RPM は各遺伝子にマップされたリード数を処理された総リード数で割り、 10^6 を乗じて算出した。ペアエンドトランスクリプトームデータの場合、プラットフォーム間の一貫性を保つため、分母は総リード数の半分（すなわちリードペアの数）とした。NumReads(fastq) は両メイト (R1+R2) の処理された総リード数を表す。RPM 計算式は以下の通りである。

CAGE/NET-CAGE の場合：

$$\text{RPM} = \left(\frac{\text{NumReads}(\text{gene})}{\text{NumReads}(\text{fastq})} \right) \times 10^6$$

ペアエンドトランスクリプトームデータの場合：

$$\text{RPM} = \left(\frac{\text{NumReads}(\text{gene})}{\text{NumReads}(\text{fastq})/2} \right) \times 10^6$$

ここで、NumReads(gene) は各遺伝子にマップされたリード数、NumReads(fastq) は処理された総リード数である。

各 *ith* 遺伝子領域内のリピート含有率について、TE 合計 (%) は以下のように算出した。ここで、 $\Sigma\text{TE}(\text{bp})$ は遺伝子領域内の全 TE アノテーションの合計長 (bp)、Gene-body(bp) は遺伝子領域長 (bp) を示す。

$$\text{TE total (\%)} = \left(\frac{\Sigma \text{TE}(\text{bp})}{\text{Gene-body}(\text{bp})} \right) \times 100$$

各リピートクラスについて、クラス別パーセンテージは以下のように算出した。ここで、Class(bp) は遺伝子領域内の対応するリピートクラスに割り当てられたアノテーションの合計長 (bp) を示す。これらの値は表にまとめ、積み上げ棒グラフとして可視化した。

$$\text{Class (\%)} = \frac{\Sigma \text{class}(\text{bp})}{\text{Gene-body}(\text{bp})} \times 100$$

ドットプロットは GENETYX 11 の HarrPlot を使い、Dot Plot Matching Num パラメータを 25 に設定して生成した。プロットは PDF としてエクスポートし、ImageMagick (v6.9.11-60; Q16 x86_64) を用いて JPEG に変換した。各手法のセクションで特に指定がない限り、すべての定量的および統計的解析は各ソフトウ

エアパッケージのデフォルト設定を用いて行った。

参考文献

- Agata, K., & Inoue, T. (2012). Survey of the differences between regenerative and non-regenerative animals. *Dev. Growth Differ.* 54, 143–152. <https://doi.org/10.1111/j.1440-169X.2011.01323.x>
- Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M, J., Kuramochi-Miyagawa, S., Nakano, T., Chien, M., Russo, J, J., Ju, J., Sheridan, R., Sander, C., Zavolan, M., & Tuschl, T. (2006). A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* 442, 203-7. <https://doi.org/10.1038/nature04916>
- Berget, S, M., Moore, C., & Sharp, P, A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci USA.* 74, 3171-5 <https://doi.org/10.1073/pnas.74.8.3171>
- Brown, T., Mishra, K., Elewa, A., Iarovenko, S., Subramanian, E., Araus, A, J., Petzold, A., Fromm, B., Friedländer, MR., Rikk, L., Suzuki, M., Suzuki, K, T., Hayashi, T., Toyoda, A., Oliveira, C, R., Osipova, E., Leigh, N, D., Yun, M, H. & Simon, A. (2025). Chromosome-scale genome assembly reveals how repeat elements shape non-coding RNA landscapes active during newt limb regeneration. *Cell Genom.* 5, 100761. <https://doi.org/10.1016/j.xgen.2025.100761>
- Chow, L, T., Gelinas, R, E., Broker, T, R., & Roberts, R, J. (1977). An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell.*12(1):1-8. [https://doi.org/10.1016/0092-8674\(77\)90180-5](https://doi.org/10.1016/0092-8674(77)90180-5)
- Corona M, Libbrecht R, Wurm Y, Riba-Grognuz O, Studer RA, Keller L. (2013). Vitellogenin underwent subfunctionalization to acquire caste and behavioral specific expression in the harvester ant *Pogonomyrmex barbatus*. *PLoS Genet.* 2013;9(8): e1003730. <https://doi.org/10.1371/journal.pgen.1003730>
- Elewa, A., Wang, H., Talavera-López, C., Joven, A., Brito, G., Kumar, A., Hameed, S, L., Penrad-Mobayed, M., Yao, Z., Zamani, N., Abbas, Y., Abdullayev, I., Sandberg, R., Grabherr, M., Andersson, B., & Simon, A. (2017). Reading and editing the *Pleurodeles waltl* genome reveals several new features of limb regeneration and genome evolution. *Nature Communi.* <https://www.nature.com/articles/s41467-017-01964-9>

Gazolla, C. B., Ludwig, A., de Moura Gama, J., & Bruschi, D. P. (2022). Evolutionary dynamics of DIRS-like and Ngarg-like retrotransposons in *Xenopus laevis* and *Xenopus tropicalis* genomes. *G3 (Bethesda)*. *12*, jkab391.

<https://doi.org/10.1093/g3journal/jkab391>

Hayashi, T., Sakamoto, K., Sakuma, T., Yokotani, N., Inoue, T., Kawaguchi, E., Agata, K., Yamamoto, T., & Takeuchi, T. (2013). Transcription activator-like effector nucleases efficiently disrupt the target gene in Iberian ribbed newts (*Pleurodeles waltl*), an experimental model animal for regeneration. *Dev. Growth Differ.* *56*, 115–121.

<https://doi.org/10.1111/dgd.12103>

Hirabayashi, S., Bhagat, S., Matsuki, Y., Takegami, Y., Uehata, T., Kanemaru, A., Itoh, M., Shirakawa, K., Takaori-Kondo, A., Takeuchi, O., Carninci, P., Katayama, S., Hayashizaki, Y., Kere, J., Kawaji, H., & Murakawa, Y. (2019). NET-CAGE characterizes the dynamics and topology of human transcribed cis-regulatory elements. *Nature Genet.* *51*, 1369–1379. <https://doi.org/10.1038/s41588-019-0485-9>

Inoue, T., Inoue, R., Tsutsumi, R., Tada, K., Urata, Y., Michibayashi, C., Takemura, S., & Agata, K. (2012). Lens regenerates by means of similar processes and timeline in adults and larvae of the newt *Cynops pyrrhogaster*. *Dev. Dyn.* *241*, 1575–1583.

<https://doi.org/10.1002/dvdy.23854>

Kimura, Y., Suzuki, M., Okumura, A., Matsunami, M., Nishide, H., Mizuno, R., Bou, K., Uno, Y., Nakada, T., Hasunuma, I., Haramoto, Y., Fukui, A., Inoue, T., Sato, Y., Yamaguchi, K., Zhang, Z., Chihara, A., Takehara, M., Shibata, Y., Kitada, M., Moreno, N., Uchiyama, I., Suzuki, Y., Takeuchi, T., Nikaido, M., Agata, K., Toyoda, A., Shigenobu, S., Hayashi, T., & Suzuki, K. T. (2025) The inbred newt genome unveils molecular mechanisms of behavior, development, and regeneration in urodele amphibians. *iScience*. *28*(10):113535. <https://doi.org/10.1016/j.isci.2025.113535>

Kurosaka, H., Takano-Yamamoto, T., Yamashiro, T., & Agata, K. (2008). Comparison of molecular and cellular events during lower jaw regeneration of newt (*Cynops pyrrhogaster*) and west African clawed frog (*Xenopus tropicalis*). *Dev. Dyn.* *237*, 354–365. <https://doi.org/10.1002/dvdy.21419>

Maki, N., Takechi, K., Sano, S., Tarui, H., Sasai, Y., & Agata, K. (2007). Rapid

- accumulation of nucleostemin in nucleolus during newt regeneration. *Dev. Dyn.* 236, 941–950. <https://doi.org/10.1002/dvdy.21027>
- Matsunami, M., Suzuki, M., Haramoto, Y., Fukui, A., Inoue, T., Yamaguchi, K., Uchiyama, I., Mori, K., Tashiro, K., Ito, Y., Takeuchi, T., Suzuki, K. T., Agata, K., Shigenobu, S., & Hayashi, T. (2019). A comprehensive reference transcriptome resource for the Iberian ribbed newt *Pleurodeles waltl*, an emerging model for developmental and regeneration biology. *DNA Res.* 26, 217–229. <https://doi.org/10.1093/dnares/dsz003>
- Mizuno, N., Agata, K., Sawada, K., Mochii, M., & Eguchi, G. (2002). Expression of crystallin genes in embryonic and regenerating newt lenses. *Dev. Growth Differ.* 44, 251–256. <https://doi.org/10.1046/j.1440-169X.2002.00639.x>
- Mitros, T., Lyons, J. B., Session, A. M., Jenkins, J., Shu, S., Kwon, T., Lane, M., Ng, C., Grammer, T. C., Khokha, M. K., Grimwood, J., Schmutz, J., Harland, R. M., & Rokhsar, D. S. (2019). A chromosome-scale genome assembly and dense genetic map for *Xenopus tropicalis*. *Dev. Biol.* 452, 8–20. <https://doi.org/10.1016/j.ydbio.2019.03.015>
- Nowoshilow S, Schloissnig S, Fei JF, Dahl A, Pang AWC, Pippel M, Winkler S, Hastie AR, Young G, Roscito JG, Falcon F, Knapp D, Powell S, Cruz A, Cao H, Habermann B, Hiller M, Tanaka EM, & Myers EW. (2018). The axolotl genome and the evolution of key tissue formation regulators. *Nature* 554, 50–55. <https://www.nature.com/articles/nature25458>
- Ochi H, Kawaguchi A, Tanouchi M, Suzuki N, Kumada T, Iwata Y, & Ogino H. (2017). Co-accumulation of cis-regulatory and coding mutations during the pseudogenization of the *Xenopus laevis* homoeologs six6.L and six6.S. *Dev Biol.* 427(1),84-92. <https://doi.org/10.1016/j.ydbio.2017.05.004>.
- Poulter, R, T, M., & Butler, M, I. (2015). Tyrosine recombinase retrotransposons and transposons. *Mobile DNA III*, p. 1271–1291. <https://doi.org/10.1128/microbiolspec.mdna3-0036-2014>
- Poulter, R, T, M., & Goodwin, T, J, D. (2005). DIRS-1 and the other tyrosine recombinase retrotransposons. *Cytogenet Genome Res.* 110:575–588.

<https://doi.org/10.1159/000084991>

Salmela H, Harwood GP, Münch D, Elsik CG, Herrero-Galán E, Vartiainen MK, Amdam GV. (2022). Nuclear translocation of vitellogenin in the honey bee (*Apis mellifera*). *Apidologie*, 53(1), 13. <https://doi.org/10.1007/s13592-022-00914-9>.

Schloissnig S, Kawaguchi A, Nowoshilow S, Falcon F, Otsuki L, Tardivo P, Timoshevskaya N, Keinath MC, Smith JJ, Voss SR, Tanaka EM. (2021): The cas9-facilitated complete genome assembly of the axolotl. *Proc Natl Acad Sci USA*. 118(15). e2017176118. <https://doi.org/10.1073/pnas.2017176118>

Sessions, S, K., & Wake, D, B. (2020). Forever young: Linking regeneration and genome size in salamanders. *Dev. Dyn.* 250(6), 768–778. <https://doi.org/10.1002/dvdy.279>

Siomi, M, C., & Kuramochi-Miyagawa, S. (2009). RNA silencing in germlines—exquisite collaboration of Argonaute proteins with small RNAs for germline survival. *Curr. Opin. Cell Biol.* 21, 426–434. <https://doi.org/10.1016/j.ceb.2009.02.003>

Subramanian, E., Elewa, A., Brito, G., Kumar, A., Segerstolpe, Å., Karampelias, C., Björklund, Å., Sandberg, R., Echeverri, K., Lui, W, O., Andersson, O., & Simon, A. (2023). A small noncoding RNA links ribosome recovery and translation control to dedifferentiation during salamander limb regeneration. *Dev Cell.* 58(6):450-460.e6. <https://doi.org/10.1016/j.devcel.2023.02.007>

Suzuki, M., Okumura, A., Chihara, A., Shibata, Y., Endo, T., Teramoto, M., Agata, K., Bronner, M, E., & Suzuki, K, T. (2024). *Fgf10* mutant newts regenerate normal hindlimbs despite severe developmental defects. *Proc Natl Acad Sci USA*. 121(11):e2314911121. <https://doi.org/10.1073/pnas.2314911121>

Suzuki, M., Hayashi, T., Inoue, T., Agata, K., Hirayama, M., Suzuki, M., Shigenobu, S., Takeuchi, T., Yamamoto, T., & Suzuki, K,T. (2018). Cas9 ribonucleoprotein complex allows direct and rapid analysis of coding and noncoding regions of target genes in *Pleurodeles waltl* development and regeneration. *Dev. Biol.* 443, 127–136. <https://doi.org/10.1016/j.ydbio.2018.09.008>

- Takeuchi, T., Matsubara, H., Minamitani, F., Satoh, Y., Tozawa, S., Moriyama, T., Maruyama, K., Suzuki, K., Shigenobu, S., Inoue, T., Tamura, K., Agata, K., and Hayashi, T. (2022). Newt Hoxa13 has an essential and predominant role in digit formation during development and regeneration. *Development* 149, dev200282. <https://doi.org/10.1242/dev.200282>
- Tamura, K., & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10,512-526. <https://doi.org/10.1093/oxfordjournals.molbev.a040023>
- Uemasu, H., Ikuta, H., Igawa, T., Suzuki, M., Kyakuno, M., Iwata, Y., Tazawa, I., Ogino, H., Satoh, Y., Takeuchi, T., Namba, N., & Hayashi, T. (2022). Cryo-injury procedure-induced cardiac regeneration shows unique gene expression profiles in the newt *Pleurodeles waltl*. *Dev Dyn.* 251(5):864-876. <https://doi.org/10.1002/dvdy.450>
- Urata, Y., Yamashita, W., Inoue, T., & Agata, K. (2018). Spatio-temporal neural stem cell behavior leads to both perfect and imperfect structural brain regeneration in adult newts. *Biol. Open* <https://doi.org/10.1242/bio.033142>
- van den Berg, J., van Ooyen, A., Mantei, N., Schamböck, A., Grosveld, G., Flavell, R. A., & Weissmann, C. (1978). Comparison of cloned rabbit and mouse beta-globin genes showing strong evolutionary divergence of two homologous pairs of introns. *Nature* 276, 37-44. <https://doi.org/10.1038/276037a0>
- Wistow, G., & Piatigorsky, J. (1988). Lens crystallins: the evolution and expression of proteins for a highly specialized tissue. *Annual Review of Biochemistry*, 57, 479-504. <https://doi.org/10.1146/annurev.bi.57.070188.002403>
- Yaguchi H, Suzuki S, Kanasaki N, Masuoka Y, Suzuki R, Suzuki RH, Hayashi Y, Shigenobu S, & Maekawa K. (2023). Evolution and functionalization of vitellogenin genes in the termite *Reticulitermes speratus*. *J Exp Zool B Mol Dev Evol.* 340(1):68-80. <https://doi.org/10.1002/jez.b.23141>.

謝辞

本研究の遂行にあたり、指導教員として大変お世話になりました阿形 清和博士
に厚く感謝申し上げます。

スキル向上としてイモリゲノム解析に携わらせて頂いた
鈴木 賢一博士及び鈴木研究室の皆様

本審査審査員として参加いただきました

三浦 正幸博士、藤森 俊彦博士および林 利憲博士

授業科目の生命科学プログレス教員としてアドバイスをくださいました

新美 輝幸博士、作田 拓博士、大野 薫博士、高橋 弘樹博士

基礎生物学研究所、オープンラボメンバーである

西村 紀子さん、楯根 一夫博士、近藤 真紀さん、諸岡 直樹さん、坂神 真理さん

大学及び大学院進学にあたり支えていただいた両親、祖父母に

感謝申し上げます。

補足

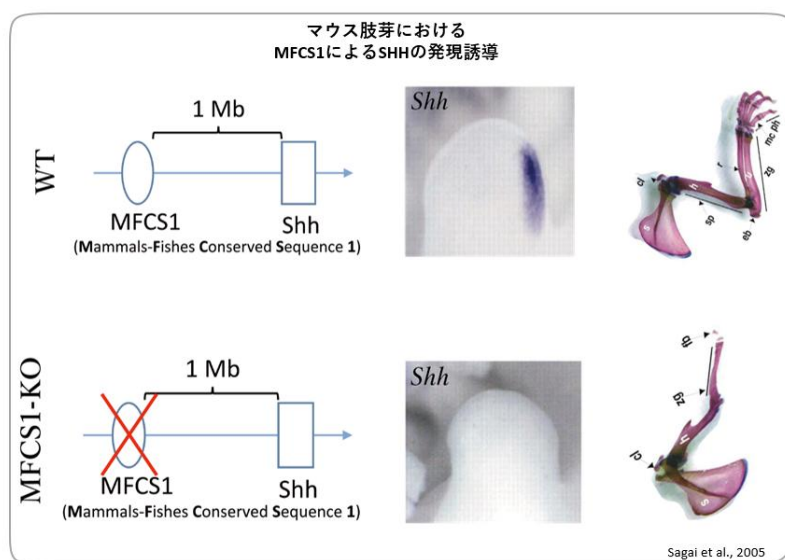
四肢特異的 *Shh* エンハンサー-MFCS1 の解析

目次

序説	55-59
結果	
MFCS1 コア領域からの明確な転写開始点は検出されない	60
MFCS1 周辺ではステージ・領域依存的な転写開始イベントが観察される	60-61
再生芽 Posterior 側で観察されるピークは convergent な 配置を示す	62
<i>Lmbr1</i> 由来の擬陽性である可能性は低い	62-63
考察	
MFCS1 コア由来の典型的 eRNA 転写は強く支持されない	64
再生芽 Posterior 側でのみ観察された convergent な配置は、典型的 divergent eRNA とは異なる現象を示唆する	64-65
「再生能差=MFCS1 単独」の単純化は難しく、制御ランドスケープ全体の再編成として捉える必要がある	65
今後の検証：転写産物の連続性とエピジェネティクス	65-66
実験材料と方法	67
参考文献	68-70

序説

四肢形成・四肢再生の分子基盤としては、発生期の四肢で *Shh* 遺伝子の時空間特異的発現とそれを制御する遠位エンハンサー MFCS1 (Mammals-Fishes Conserved Sequence 1: 別名 ZRS) が重要である。*Shh* 遺伝子は、肢芽の後部側周縁に位置する極性化活性帯分 (ZPA: Zone of Polarizing Activity) で発現し、脊椎動物の四肢発生において、指を含む前後軸パターンの形成と四肢の伸展に重要な役割を担っていることが予測されていた (Todt & Fallon, 1987; Riddle et al., 1993; Niswander et al., 1994; Haramis et al., 1995; Sagai et al., 2004)。実際に、*Shh* コード領域から 1 Mb 離れた非コード領域に保存されている MFCS1 領域をマウスのゲノムから遺伝子ターゲティングの手法で除くと、肢芽での *Shh* 遺伝子の特異的な発現が消失し、指をもたないスパイク状の四肢が発生することが報告されていた(補足図 1; Sagai et al., 2005, Masuya et al., 2007, Sagai et al., 2009)。

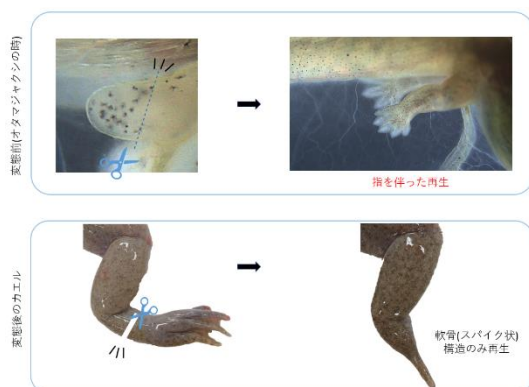


補足図 1 マウスで報告された MFCS1 配列の欠損による *Shh* の発現消失と四肢の形態変化

MFCS1 が KO されたマウスでは肢芽 ZPA での *Shh* 遺伝子の発現が消失し、指が形成されずにスパイク状の形の四肢が発生する。

カエルでは、変態前は後肢を伸長中に切断しても、指を伴った肢を再生できるが、変態後に切断するとスパイク状の後肢が再生すること、変態後のカエル

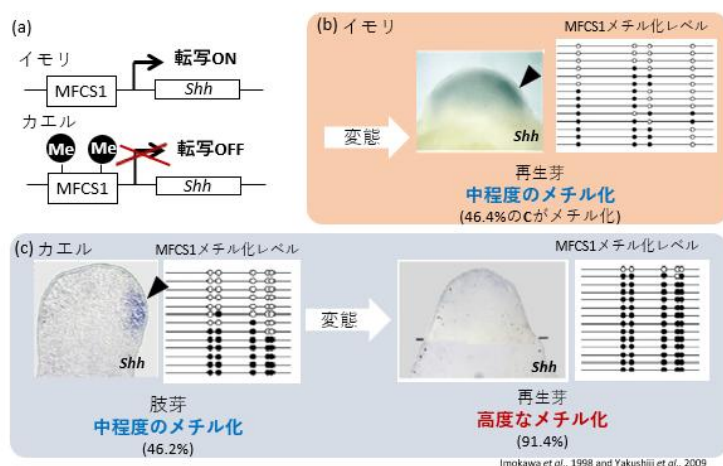
の再生芽では *Shh* 遺伝子の発現がないことが古くから知られていた(補足図 2、Yokoyama et al., 2011)。マウスでの MFCS1 の KO マウスがスパイク状の四肢を形成したことから、カエルでは変態後に MFCS1 の機能の低下が原因なのではないかとの推察が行われた。



補足図 2 カエルでは変態前後で再生能が変化する

変態前の肢芽の切断は指を伴った再生ができるのに対して、変態後の四肢を切断すると、MFCS1 の KO マウスと良く似た指のないスパイク構造のみを再生する。

そこで、東北大学のグループがカエルのゲノムの MFCS1 領域を bisulfite 法で解析したところ、MFCS1 領域の DNA のメチル化の程度と再生能との相関があることを報告し(補足図 3)、カエルでは変態後に MFCS1 がエピジェネティックな修飾を受けることで再生能を失うことが示唆された(Yakushiji et al., 2007)



補足図 3 カエル肢芽、成体再生芽とイモリ再生芽の *Shh* 発現と MFCS1 のメチル化状態

イモリ再生芽とカエル肢芽ではメチル化レベルが高くなく、*Shh* 遺伝子が発現しているのに対し、成体カエルの四肢再生芽ではメチル化レベルの高く、*Shh* 遺伝子が見られてない。

近年、遠位エンハンサーの役割として、遠位エンハンサーから non-coding RNA が転写され、それが、エンハンサーとプロモーターとのループ構造の形成

を補助することを示唆する論文が報告された(補足図4、Matharu et al., 2015)。そして、このエンハンサー領域から転写される non-coding RNA はエンハンサーRNA (eRNA)と命名された。

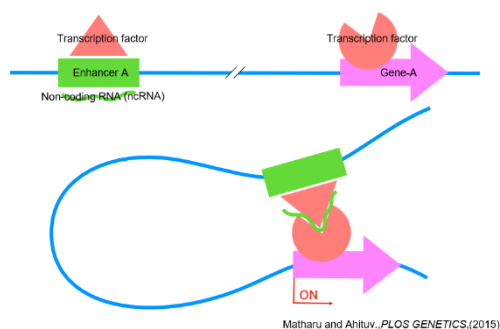
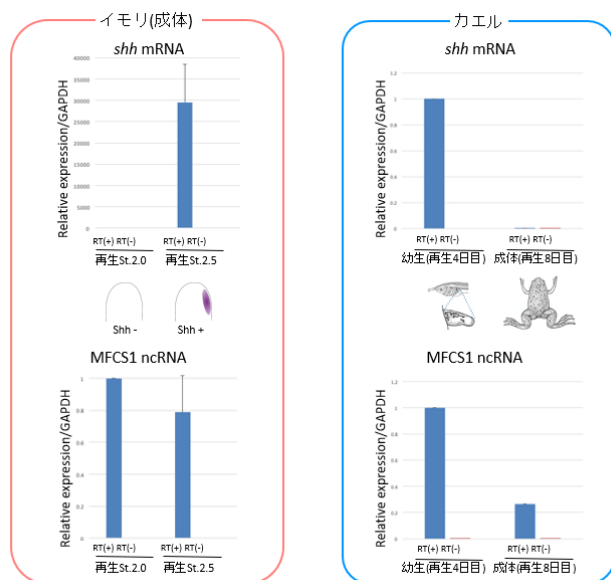


図4 提唱されたエンハンサーRNAのイメージ図

エンハンサー部から転写されたeRNAを介してエンハンサー領域とプロモーター領域が接近することで、標的遺伝子が活性化されると考えられている(Matharu & Ahituvらの論文から改変した図を作ったもの)。

そこで所属研究室ではMFCS1もeRNAとして転写されることで5Mbに離れた*Shh*遺伝子の活性化が行われている可能性を検証するために、MFCS1からの転写産物のqRT-PCRによる検出が行われた。



補足図5 MFCS1からのncRNAの発現と*Shh* mRNAとの相関

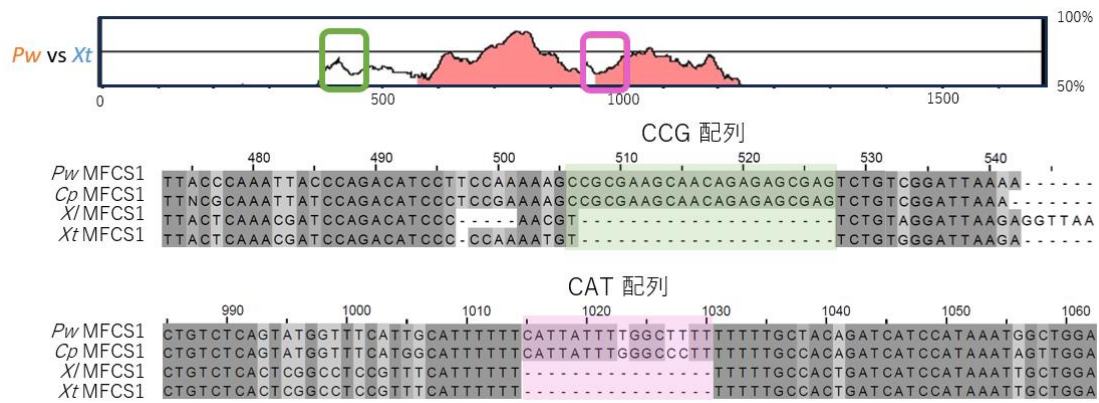
イモリとカエルのそれぞれのMFCS1領域のPCRプライマーを用いてRNA転写産物の定量PCRをしたところ、*shh*遺伝子発現に先行する形でncRNAが検出された(学習院大学・生物学科・角田瑛、修士論文より転載)

特にイモリ再生芽においてその結果、*Shh*遺伝子が発現する直前のステージでMFCS1由来のRNAが検出された。これによりMFCS1はeRNAとして機能している可能性が示唆され、エ

ンハンサー領域のエピジェネティックな転写と再生能との相関が疑われた (補足図 5)。

そこで本研究では、MFCS1 が enhancer RNA として転写され得るという可能性に注目し、転写開始点 (TSS) の検出を目的として NET-CAGE を行なった。成体肝組織に加え、四肢発生期の肢芽および成体再生芽を対象に解析した結果、肝組織では高発現遺伝子 (例: *Fth*, *Alb*) の TSS を同定でき、肢芽・再生芽でも *Shh* の TSS を検出できたことから、TSS 解析に耐えるデータが得られた。イベリアトゲイモリとネッタイツメガエルの MFCS1 領域の比較から、2 種のイモリに特異的に保存されていた CCG 配列部位(補足図 6 緑部分)と CAT 配列(補足図 6 赤部分)に着目した。これらの配列から eRNA の転写が起きていることを期待した。CCG 配列はプロモーター近傍で転写される双方向性 non-coding RNA (panc RNA: promoter-associated non-coding RNA の略)の転写開始点で保存されている配列に近いことから、転写開始点の候補として期待された。

また、MFCS1 領域の両外側にも 2 種のイモリで塩基配列が高く保存されている領域も見出されており、そこからの転写についても NET-CAGE 解析で検討を行った。



補足図 6. VISTA 解析による MFC1 近傍で、イモリとカエルで保存されている領域
 約 1.6kb の MFC1 領域に全般において高い保存性が見いだされると同時にイモリ科に固有の配列が二箇所に見いだされたので、それぞれに CCG 配列(緑で表示)、CAT 配列(赤で表示)という名称をつけた。

結果

MFCS1 コア領域からの明確な転写開始点は検出されない

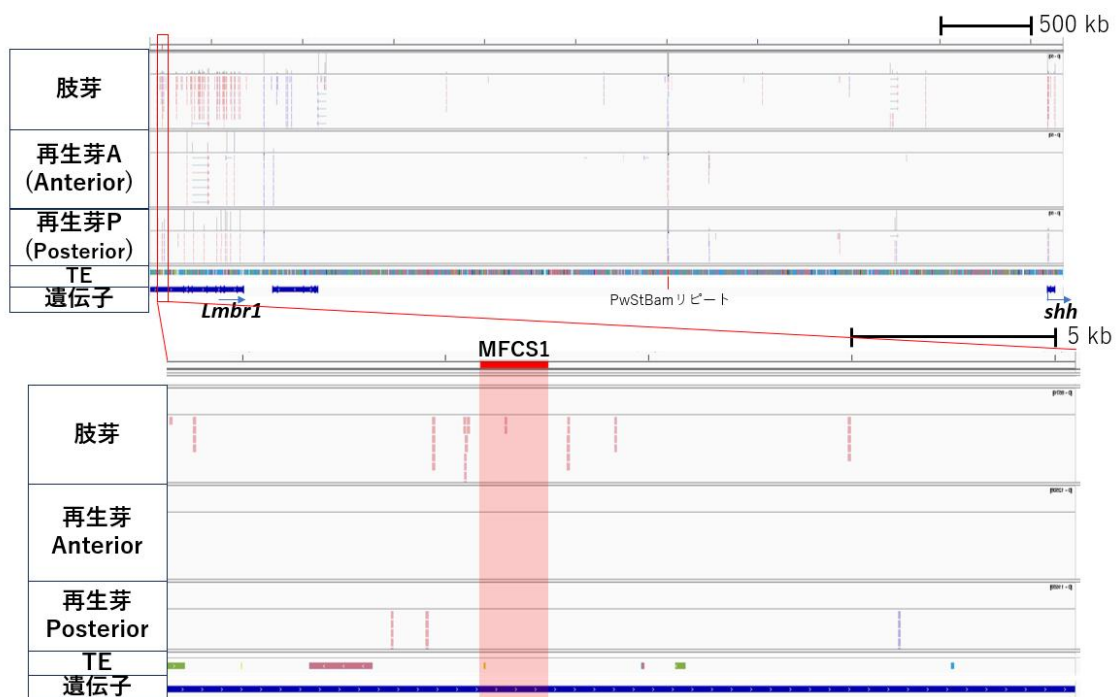
MFCS1 が enhancer RNA (eRNA) として転写される可能性を検証するため、イモリの肢芽および再生芽に対して NET-CAGE を実施し、MFCS1 転写開始点 (TSS) の検出を試みた。再生芽は ZPA を含み *Shh*/MFCS1 活性が想定される Posterior 側と、ZPA を欠く Anterior 側に分けて解析した。その結果、MFCS1 として定義した 1.6 kb コア領域内からの NET-CAGE リードは、いずれのサンプルでも明確には検出されなかった (補足図 7a、赤枠部分)。肢芽でのみ 1 リードが検出されたが、再現性や局在の点から、TSS として使用されていることを支持するには弱い結果であった。

MFCS1 周辺ではステージ・領域依存的な転写開始イベントが観察される

次に、MFCS1 コア領域の周辺領域を拡張して解析したところ、肢芽、再生芽 Anterior 側、再生芽 Posterior 側で、転写開始イベントのパターンが大きく異なることが分かった (補足図 7a)。

発生途中の肢芽では、*Shh* 遺伝子方向に向かう複数の NET-CAGE ピークが観察された。一方、再生芽 Anterior 側では MFCS1 近傍でピークは全く認められなかった。ZPA を欠く領域では *Shh*/MFCS1 活性が低いという既報の知見と合致し

た。対照的に、再生芽 Posterior 側では MFCS1 を挟む形でピークが検出され、
 枝芽および再生芽 Anterior 側には見られない特徴的な配置を示した。唯一の共通点としては、*Shh* 遺伝子を発現している枝芽と再生芽 Posterior 側とでだけ、
 MFCS1 配列を含む non-coding RNA を転写する転写開始産物が見つかった点で
 ある。



補足図 7 イモリ枝芽および再生芽の NET-CAGE のマッピング結果

- (a) MFCS1 1.6 kb とその周辺のマッピング結果。青が *Shh* 遺伝子方向に向かう(+)ストランド、赤が逆方向の転写を示す(-)ストランドとなっている。
- (b) *Shh* から MFCS1 までのゲノム領域のマッピング結果。右端の赤い枠は(a)の領域を示す。*Shh* から MFCS1 までは約 4.9 Mb の距離がある。

再生芽 Posterior 側で観察されるピークは convergent な配置を示す

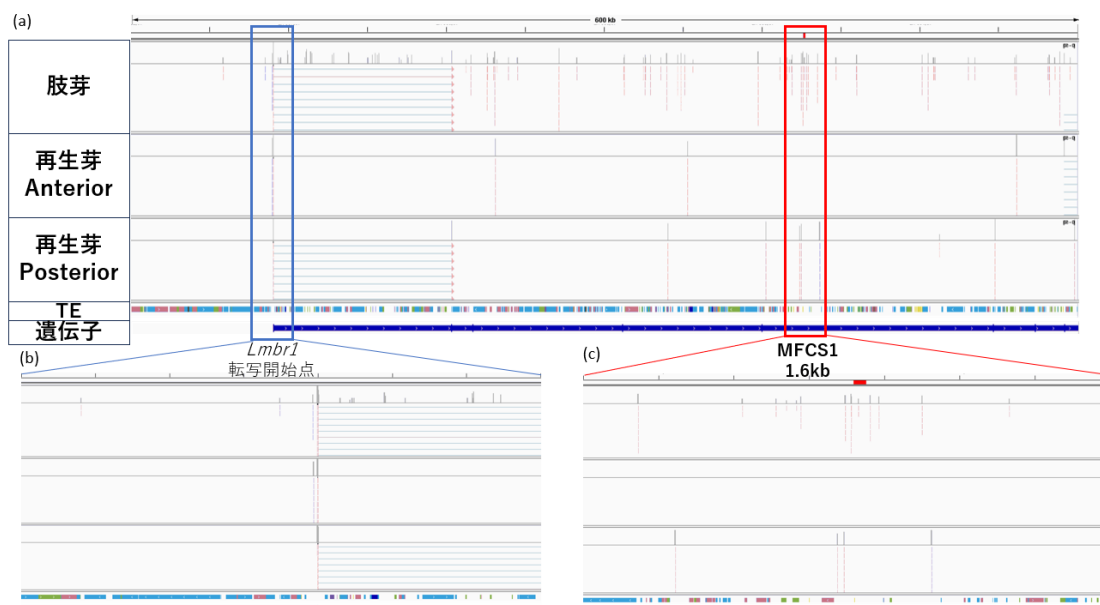
IGV により strand-separate 表示で確認したところ、再生芽 Posterior 側では、(+)ストランド由来のリードが MFCS1 上流側に二箇所、(-)ストランド由来のリードが MFCS1 下流側に一箇所マッピングされ、MFCS1 を挟んで互いに内向き (convergent) に配置されたピークとして観察された (補足図 7a)。ただ、この配置は、活性化エンハンサーで典型的とされる外向きの双方向 (divergent) 転写とは異なる。

また、MFCS1 配列から *Shh* 遺伝子まで約 5 Mb 近く離れており (補足図 7b)、その間の領域は TE によってほとんど埋め尽くされていた。この領域から TE 配列と重なるリードを除外すると、この約 5 Mb にわたる領域においてほとんど noncoding RNA と考えられる配列が転写されていなかった。

Lmbr1 由来の擬陽性である可能性は低い

過去に qRT-PCR により MFCS1 由来 RNA の存在が示唆されていたが、その際には MFCS1 が *Lmbr1* 遺伝子のイントロン内に存在していることから、pre-mRNA を検出している可能性が否定できなかった(所属研究室データ)。今回の NET-CAGE データでは、*Lmbr1* の主要な TSS ピークと MFCS1 近傍のピークとの間に連続したつながりは認められず(補足図 8 青枠)、MFCS1 近傍のピークが単純に *Lmbr1* の未スプライス転写産物に由来する可能性は低いと考えられた。

さらに、TE アノテーション上、MFCS1 およびその周辺は TE 侵入が乏しい保存領域であり(補足図 8a,b)、検出されたピークは TE と重ならない位置に局在していた。以上より、再生芽 Posterior 側で観測された MFCS1 周辺の転写開始イベントは、少なくとも TE 由来の非特異的転写ではなく、MFCS1 を含む制御ランドスケープに関連した転写開始である可能性が示唆された。



補足図 8 *Lmbr1* 転写開始点と MFCS1 周辺の俯瞰図と拡大図

- (a) MFCS1 から *Lmbr1* 転写開始点までの俯瞰図
- (b) MFCS1 周辺の NET-CAGE リード。再生芽では TE のない場所からの転写がある。
- (c) *Lmbr1* 転写開始点周辺の NET-CAGE リード。転写開始点からの高いピークがある。

考察

MFCS1 コア由来の典型的 eRNA 転写は強く支持されない

本研究の当初の目的は、MFCS1 が eRNA として転写され、*Shh* との長距離ループ形成（あるいはその維持）を補助する可能性を NET-CAGE により検出することであった。しかし、MFCS1 と定義した 1.6 kb 領域内部からの明確な転写開始点は、枝芽・再生芽いずれにおいても検出されなかった。したがって、少なくとも本解析条件下では、MFCS1 内部に安定した TSS を持つ「MFCS1 そのものの eRNA」を検出できなかったとは言えないが、次のステップとして、イモリの MFCS1 をカエルにゲノム編集で挿入しても *Shh* 遺伝子の発現を誘導できるようになるとは考えにくい。このため今後、実験デザインの再考が必要であろう。

再生芽 Posterior 側でのみ観察された convergent な配置は、典型的 divergent eRNA とは異なる現象を示唆する

一方で、MFCS1 周辺では枝芽と再生芽 (Posterior) で転写開始パターンが異なり、とくに再生芽 Posterior 側では MFCS1 を挟む内向き (convergent) 配置のストランド反対向きピークが観測された。活性化エンハンサーでよく報告される双方向性は、単一領域から両端へ外向きに転写が立ち上がる divergent 配置として説明されることが多い(Uesaka et al., 2014)。今回の配置はそれとは異なるた

め、「典型的な bidirectional eRNA」と断定するのは適切ではない。むしろ、再生芽 ZPA において MFCS1 周辺の局所的な転写開始が特異的に誘導される、別種の転写イベント (enhancer-associated initiation) として捉える方が妥当かと考えられる。

「再生能差 = MFCS1 単独」の単純化は難しく、制御ランドスケープ全体の再編成として捉える必要がある

本研究では、イモリと成体カエルの四肢再生能の差の説明要因として期待した MFCS1 内部の明瞭な転写を直接には支持できなかった。しかし、再生芽 Posterior 側でのみ MFCS1 周辺に特異な転写開始イベントが観察されたことは、再生芽 ZPA において *Shh* 制御ランドスケープが発生期とは異なる様式で再編成される可能性を示している。したがって、再生能の差を単一エンハンサー配列の有無や単一因子の活性に還元するのではなく、MFCS1 を含む広域制御領域のクロマチン状態や転写開始の組み合わせとして理解する必要があると考えられた。

今後の検証：転写産物の連続性とエピジェネティクス

今後、*Shh* 遺伝子の転写イベントの実体を明らかにするには、(i) MFCS1 周辺ピークからの RNA-seq カバレッジの連続性 (短命な eRNA 様か、より長い転写

単位か)、(ii) H3K27ac/H3K4me1 や ATAC-seq による活性化エンハンサー標識との一致、(iii) 5'RACE 等による TSS の独立検証、が有効と考えられた。これらを統合することで、再生芽 ZPA で誘導される MFCS1 周辺転写物が *Shh* 遺伝子の発現制御にどのように関与し得るか、あるいは関与しない副産物なのかを判別できると考えられる。

実験材料と手法

第四章で記述した実験材料と手法を踏襲した。枝芽および再生芽からの RNA 抽出の方法について追記する

NET-CAGE (新生 RNA) および CAGE (全 RNA) のための RNA 抽出

イモリ枝芽は発生ステージ 34 を目標にした。枝芽が伸長し始めると同時期に鰓の分岐が始まることを見出した。これを参考に氷で冷やすことで発生を遅らせ枝芽を回収した。回収した枝芽は 500 μ l の 1×PBS(-)に 25 μ M α -アマニチン、1x complete Protease Inhibitor Cocktail、20 U SUPERase IN RNase Inhibitor となるように調整した溶液に入れ、氷上で保存した。その後肝組織破碎と同様に処理した。再生芽は血管再生ステージ 2 (血管が指様に伸長する前) を参考に Anterior 側と Posterior 側に 2 分して回収した。枝芽回収時と同様に 500 μ l の 1×PBS(-)に 25 μ M α -アマニチン、1x complete Protease Inhibitor Cocktail、20 U SUPERase IN RNase Inhibitor となるように調整した溶液に入れ、氷上で保存した。その後肝臓組織破碎と同様に処理した。

参考文献

- Agata, K., & Inoue, T. (2012). Survey of the differences between regenerative and non-regenerative animals. *Dev. Growth Differ.* 54, 143–152. <https://doi.org/10.1111/j.1440-169X.2011.01323.x>
- Chow, L. T., Gelinas, R. E., Broker, T. R., & Roberts, R. J. (1977). An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell*.12(1):1-8. [https://doi.org/10.1016/0092-8674\(77\)90180-5](https://doi.org/10.1016/0092-8674(77)90180-5)
- Haramis, A. G., Brown, J. M., & Zeller, R. (1995). The limb deformity mutation disrupts the SHH/FGF-4 feedback loop and regulation of 5' HoxD genes during limb pattern formation. *Development*, 121(12), 4237–4245. <https://doi.org/10.1242/dev.121.12.4237>
- Hirabayashi, S., Bhagat, S., Matsuki, Y., Takegami, Y., Uehata, T., Kanemaru, A., Itoh, M., Shirakawa, K., Takaori-Kondo, A., Takeuchi, O., Carninci, P., Katayama, S., Hayashizaki, Y., Kere, J., Kawaji, H., & Murakawa, Y. (2019). NET-CAGE characterizes the dynamics and topology of human transcribed cis-regulatory elements. *Nature Genet.* 51, 1369–1379. <https://doi.org/10.1038/s41588-019-0485-9>
- Imokawa, Y., & Yoshizato, K. (1997). Expression of Sonic hedgehog gene in regenerating newt limb blastemas recapitulates that in developing limb buds. *Proceedings of the National Academy of Sciences of the United States of America*, 94(17), 9159–9164. <https://doi.org/10.1073/pnas.94.17.9159>
- Kimura, Y., Suzuki, M., Okumura, A., Matsunami, M., Nishide, H., Mizuno, R., Bou, K., Uno, Y., Nakada, T., Hasunuma, I., Haramoto, Y., Fukui, A., Inoue, T., Sato, Y., Yamaguchi, K., Zhang, Z., Chihara, A., Takehara, M., Shibata, Y., Kitada, M., Moreno, N., Uchiyama, I., Suzuki, Y., Takeuchi, T., Nikaido, M., Agata, K., Toyoda, A., Shigenobu, S., Hayashi, T., & Suzuki, K. T. (2025) The inbred newt genome unveils molecular mechanisms of behavior, development, and regeneration in urodele amphibians. *iScience*. 28(10):113535. <https://doi.org/10.1016/j.isci.2025.113535>
- Masuya, H., Sezutsu, H., Sakuraba, Y., Sagai, T., Hosoya, M., Kaneda, H., Miura, I., Kobayashi, K., Sumiyama, K., Shimizu, A., Nagano, J., Yokoyama, H., Kaneko, S., Sakurai, N., Okagaki, Y., Noda, T., Wakana, S., Gondo, Y., & Shiroishi, T. (2007). A

- series of ENU-induced single-base substitutions in a long-range cis-element altering Sonic hedgehog expression in the developing mouse limb bud. *Genomics*, 89(2), 207–214. <https://doi.org/10.1016/j.ygeno.2006.09.005>
- Navneet, M., & Nadav, A. (2015). Minor loops in major folds: Enhancer–promoter looping, chromatin restructuring, and their association with transcriptional regulation and disease. *PLoS Genetics*, 11(12), e1005640. <https://doi.org/10.1371/journal.pgen.1005640>
- Niswander, L., Jeffrey, S., Martin, G. R., & Tickle, C. (1994). A positive feedback loop coordinates growth and patterning in the vertebrate limb. *Nature*, 371(6498), 609–612. <https://doi.org/10.1038/371609a0>
- Riddle, R. D., Johnson, R. L., Laufer, E., & Tabin, C. (1993). Sonic hedgehog mediates the polarizing activity of the ZPA. *Cell*, 75(7), 1401–1416. [https://doi.org/10.1016/0092-8674\(93\)90626-2\[5\]](https://doi.org/10.1016/0092-8674(93)90626-2[5])
- Sagai, T., Masuya, H., Tamura, M., Shimizu, K., Yada, Y., Wakana, S., Gondo, Y., & Shiroishi, T. (2004). Phylogenetic conservation of a limb-specific, cis-acting regulator of Sonic hedgehog (Shh). *Mammalian Genome*, 15(1), 23–34. <https://doi.org/10.1007/s00335-003-2317-5>
- Sagai, T., Hosoya, M., Mizushina, Y., Tamura, M., & Shiroishi, T. (2005). Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. *Development*, 132(4), 797–803. <https://doi.org/10.1242/dev.01613>
- Schloissnig S, Kawaguchi A, Nowoshilow S, Falcon F, Otsuki L, Tardivo P, Timoshevskaya N, Keinath MC, Smith JJ, Voss SR, Tanaka EM. (2021): The cas9-facilitated complete genome assembly of the axolotl. *Proc Natl Acad Sci USA*. 118(15). e2017176118. <https://doi.org/10.1073/pnas.2017176118>
- Suzuki, M., Okumura, A., Chihara, A., Shibata, Y., Endo, T., Teramoto, M., Agata, K., Bronner, M. E., & Suzuki, K. T. (2024). *Fgf10* mutant newts regenerate normal hindlimbs despite severe developmental defects. *Proc Natl Acad Sci USA*. 121(11):e2314911121. <https://doi.org/10.1073/pnas.2314911121>

- Todt, W. L., & Fallon, J. F. (1987). Posterior apical ectodermal ridge removal in the chick wing bud triggers a series of events resulting in defective anterior pattern formation. *Development*, *101*(3), 501–515. <https://doi.org/10.1242/dev.101.3.501>
- Tsutsumi, R., Inoue, T., Yamada, S., & Agata, K. (2015). Reintegration of the regenerated and the remaining tissues during joint regeneration in the newt *Cynops pyrrhogaster*. *Regeneration* *2*, 26–36. <https://doi.org/10.1002/reg2.28>
- Tsutsumi, R., Yamada, S., & Agata, K. (2016). Functional joint regeneration is achieved using reintegration mechanism in *Xenopus laevis*. *Regeneration* *3*, 26–38. <https://doi.org/10.1002/reg2.49>
- Uesaka, M., Nishimura, O., Go, Y., Nakashima, K., Agata, K., & Imamura, T. (2014). Bidirectional promoters are the major source of gene activation-associated non-coding RNAs in mammals. *BMC Genomics*, *15*, 35. <https://doi.org/10.1186/1471-2164-15-35>
- Yakushiji, N., Suzuki, M., Satoh, A., Sagai, T., Shiroishi, T., Kobayashi, H., Sasaki, H., Ide, H., & Tamura, K. (2007). Correlation between Shh expression and DNA methylation status of the limb-specific Shh enhancer region during limb regeneration in amphibians. *Developmental Biology*, *312*(1), 171–182. <https://doi.org/10.1016/j.ydbio.2007.09.022>
- Yokoyama H, Maruoka T, Ochi H, Aruga A, Ohgo S, Ogino H, Tamura K. (2011). Different requirement for Wnt/ β -catenin signaling in limb regeneration of larval and adult *Xenopus*. *PLoS One*. *6*(7):e21721. <https://doi.org/10.1371/journal.pone.0021721>.