

機関番号：12702

研究種目：若手研究（A）

研究期間：2007～2010

課題番号：19681020

研究課題名（和文） 遺伝子重複による新規機能獲得遺伝子の同定

研究課題名（英文） Detecting neofunctionalized genes in a genome

研究代表者 印南 秀樹（INNAN HIDEKI）

総合研究大学院大学・先導科学研究科・准教授

研究者番号：90444140

研究成果の概要（和文）：本研究課題では、遺伝子重複によって新しい遺伝子が創りだされる進化メカニズムをゲノムスケールで解明することにより、ゲノム進化とそれに伴う表現型進化を理解することを究極のゴールとする。ゲノムが遺伝子重複により新規機能遺伝子を獲得するときの、進化メカニズムの一般的法則を理解することにある。そして、その結果をもとに、新規機能遺伝子をゲノム中から探し出し、それらに対して集団遺伝及び分子進化解析を行うことにより、ゲノム進化のメカニズムの解明を目指す。平成19年度、20年度では、重複した遺伝子が集団中に固定していく過程などに関する集団遺伝の理論を確立した。つづいて、平成21年度、22年度では、その応用に重点を置いた。応用範囲は、ハエ、イネ、魚に及ぶ。さらに、新規機能獲得遺伝子を同定するアルゴリズムの試運転をハエと酵母を用いて行った。とくに魚類と酵母で新規の結果が得られた。これらについて、新規機能獲得の進化的解析を行った。それによると、新規機能獲得遺伝子は、多くの近縁種にも共通にみられ、獲得イベントがそれらの種分化よりも古く、その後それぞれの種の進化のプロセスで新規機能遺伝子が自然選択の力で守られてきたことを示唆する。このように、生物が進化する上で、遺伝子重複によって新規機能獲得遺伝子を獲得することが重要な役目をはたすことが確認出来た。そして、その痕跡はゲノム中にはっきり残っており、それは本研究で確立した理論によって効果的に検出できることが分った。

研究成果の概要（英文）：This project was aimed to theoretically understand the evolutionary mechanism by which a genome acquires novel functional genes through gene duplication. The theoretical result should contribute to the developing algorithms to detect new genes from genomes. In the first half of the funded period (H19 and H20), I successfully developed theories on the pattern of DNA polymorphism and divergence in duplicated genes. This directly resulted in a new algorithm to scan a genome for evidence of newly arisen functional genes. The algorithm is simple enough to apply to any genome, from yeast to human. In the second half (H21 and H22), by using this algorithm, we identified a number of such new genes in yeasts and fruit flies, which gave insights into what kind of genes contributed the recent genome evolution of these species.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2007年度	4,300,000	1,290,000	5,590,000
2008年度	3,900,000	1,170,000	5,070,000
2009年度	3,800,000	1,140,000	4,940,000
2010年度	3,900,000	1,170,000	5,070,000
年度			
総計	15,900,000	4,770,000	2,067,000

研究分野：基礎ゲノム科学

科研費の分科・細目：ゲノム生物学

キーワード：ゲノム / シロイヌナズナ / 分子進化 / 進化 / 遺伝子変換 / 遺伝子重複 / 酵母 / 集団遺伝学

### 1. 研究開始当初の背景

生命体が進化的タイムスケールで環境に応答してきた背景には、新規遺伝子、それに伴う新規表現型の獲得があった。遺伝子重複は新規遺伝子の源であり、ゲノム進化におけるその重要性は言うまでもない(Ohno 1970)。本研究課題では、遺伝子重複によって新しい遺伝子が創りだされる進化メカニズムをゲノムスケールで解明することにより、ゲノム進化とそれに伴う表現型進化を理解することを究極のゴールとする。

重複遺伝子の進化に特徴的なものの一つに、「協調進化」がある。すなわち、重複した遺伝子コピーが遺伝子変換や不等交叉というメカニズムによって、DNA 配列を互いに交換し合いながら共進化するのである。この現象は30年以上も前にリボゾーム遺伝子において発見された。しかしながら、学会は「協調進化はリボゾーム遺伝子などの特殊な遺伝子で起こるもの」という固定観念に、つい最近まで支配されてきた。弊害として、重複遺伝子に関するゲノムデータ解析の99%以上が、この固定観念を前提として行われていることが挙げられる。

申請者の研究室では、この固定観念を壊すことに成功した。もともと申請者は、重複遺伝子の基礎的な集団遺伝理論の開発をしていた(Innan 2002,2003, Teshima and Innan 2004)。そして、これらの理論的結果を酵母のゲノムデータに適用することによって、あらゆる重複遺伝子が協調進化を経験していることを証明した(Gao and Innan 2004)。同時に、それまで頻繁に行われていた、協調進化を無視したゲノムデータ解析が、いかに間違った結論を導くものかということを示した。例えば、協調進化を無視した解析から得られた遺伝子重複率の推定は、真の値から100倍以上も高いものになる。これらの研究は、重複遺伝子の進化のメカニズムを理解する上では、非常に衝撃的なものであり、重複コピーの共進化をふまえた理論とデータ解析ツール開発が急務であることを意味する。そして、このような問題解決が、申請者の現在進行中のプロジェクトの中心になってきている。

その中で、本申請課題は、遺伝子重複による新規機能遺伝子獲得にテーマを絞る。一般的に、重複によってできた“冗長な”遺伝

子は、機能的制約を受けにくいいため比較的自由に進化し、新規機能を獲得することがしばしばある。このプロセスには、もちろん自然選択の力が大きく関与する。本申請課題では、この遺伝子重複による新規機能遺伝子獲得の進化メカニズムの解明と、ゲノム進化におけるその役割の理解を目指す。一般的に遺伝子重複は、塩基レベルの突然変異よりも表現型に対する効果は大きく、従って自然選択の影響を受けやすい。遺伝子重複による適応進化の例としては、ヒトの赤緑色覚遺伝子あげられる。ヒト近縁種では、赤と緑を認識するタンパクを重複した別々の遺伝子で作りに出しているが、新世界ザルではこの遺伝子重複が起こっていないため、赤緑の認識が困難である。この事実は、遺伝子重複によって、2色色覚システムからより高等な3色色覚システムへ適応進化したことを示唆する。このような、重複した遺伝子が新規機能を獲得するタイプの適応進化の例は多く知られている。一方で、遺伝子重複は進化的に負の影響も生み出した。すなわち、重複遺伝子が遺伝病の原因になるケースは多い。ヒトの赤緑色覚遺伝子も例外ではなく、この重複遺伝子における突然変異が色盲色弱の原因になる。このように、遺伝子重複とそれに働く正と負の自然選択は、ゲノム進化の大きな原動力の一つとなっている。そして、自然選択のモード(方向性と強さ)は、時とともに変化するものである。例えば、我々の祖先がアフリカから全世界に移住した際、新しい土地と気候、それに伴う食べ物などの変化などに遺伝的に対応してきた。すなわち、ここ数万年という進化的には非常に短い時間に、劇的な環境変化を経験し、それに合わせて、種レベルでゲノム構成を適応させて来た過程に、遺伝子重複による新規機能遺伝子獲得も重要な役割を担ったということである。

ここでは、わかりやすい例としてヒトの赤緑色覚遺伝子を用いたが、本申請課題のメインテーマは、ヒトに限ったものではない。すべての生物種に共通の進化の法則を見つけ出すことにある。そのために、ゲノムデータ解析と同時に、集団遺伝及び分子進化の理論的研究も行う。このようなモデルベースの解析を行うことにより、記述的になりがちなゲノムデータ解析研究とは一線を画し、進化のメカニズム解明にまで切り込む。そして、

重複遺伝子のゲノム進化に対する貢献を理解することを通して、我々ヒトとそれを取り巻く他の生物の歴史を広い意味で理解したい。

## 2. 研究の目的

本研究の目的は、ゲノムが遺伝子重複により新規機能遺伝子を獲得するときの、進化メカニズムの一般的法則を理解することにある。そして、その結果をもとに、新規機能遺伝子をゲノム中から探し出し、それらに対して集団遺伝及び分子進化解析を行うことにより、ゲノム進化のメカニズムの解明を目指す。

本研究課題では、理論及びゲノムデータ解析の二通りのアプローチを用いる。理論的研究としては、重複遺伝子の集団遺伝及び分子進化解析の確立と発展をテーマにする。具体的には、遺伝子が重複し、それが集団中に固定する過程の集団遺伝学解析に始まり、さらに重複遺伝子で起こった有用な突然変異が、どのように自然選択のプレッシャーのもと、ゲノム進化に貢献するかを考える。ここでは、遺伝子重複による正の効果（新規機能獲得）と負のコスト（例えば、遺伝病）を同時に考える必要がある。この正と負の効果のバランスが、環境の変化と応答しながらゲノムは進化する。これらを総合的に理解することにより、広い意味での生命体の進化をゲノム進化の視点から考える。

上の理論研究は、より効果的なデータ解析ツールのデザインに直接貢献する。具体的には、強い正の効果によって新規機能を獲得した重複遺伝子を、ゲノム中から探し出すアルゴリズムを開発する。その為には、そのような遺伝子が進化の過程でどのような特徴的なプロセスを経験するかを知る必要がある。ここで、上記の理論研究と密接にリンクしてくる。アルゴリズムの開発段階で、それが未完成の状態であっても、積極的にゲノムデータに試験適用していく。そうすることにより、アルゴリズム開発段階では発見しなかった問題点の認識、そしてその部分の改良につなげる。

最終的には、精巧なアルゴリズムの開発、そして、多くのゲノムデータへの適用を行う。現在、超多量のゲノムデータが公開されており、これらを最大限に利用する。そこから、ゲノム進化における遺伝子重複の重要性を理解し、さらには遺伝子重複により新規機能遺伝子を獲得する進化プロセスの一般法則を見つけ出す。

## 3. 研究の方法

本研究課題は、理論およびゲノムデータ解析の二通りのアプローチで、遺伝子重複によるゲノム進化のメカニズムを解明することを目的とする。特に、遺伝子重複による新規機能遺伝子獲得にテーマを絞る。これは比較的新しい分野であるため、**hypothesis generating** 的なスタイルで研究を行う。そのため、常に二つのアプローチを同時進行させる。すなわち、常に新しいデータに細心の注意を払いながら理論開発を行い、そのプロセスの途中の段階において、データと理論を見比べることにより、理論モデルがデータにあった現実的なものであるかを随時チェックする。そして、その結果は更なるデータ解析に還元される。また、データベースにはないデータが必要な場合は（例えば、種内多型データ）、塩基配列決定等の実験を行う。これら三つが、相互に作用しながら研究は遂行される。

## 4. 研究成果

遺伝子が重複すると、ゲノムは同一の遺伝子を二つもつことになる。この場合、もとのひとつがその役目を果たしている限り、二つ目の果たす役割は少ない。従って、多くの場合、二つ目は、事実上なくなってしまう。これは、その二つ目の遺伝子に突然変異が蓄積し、意味のあるタンパクをコード出来なくなってしまうことを意味する。突然変異のほとんどは、基本的に有害であるからである。しかしながら、非常に稀ではあるが、二つ目のコピーに有用な突然変異が起こることがある。例えば、もとの機能変えることによって新しい機能を与える突然変異（新規機能突然変異）などが相当する。このような突然変異が起こると、二つ目のコピーに存在意義が生まれ、ゲノム中に維持されることになる。この過程における DNA 塩基配列の進化を、理論的に研究した (Teshima Innan 2008)。

その結果、新規機能突然変異の周辺だけ特異的に塩基配列の分化度が高くなることが分った。この現象の背景には、遺伝子変換という重複遺伝子間の塩基配列を均一化するメカニズムがある。実際に、2コピー間の塩基配列の進化をコンピュータシミュレーションによって再現し、塩基配列を比較してみると、新規機能突然変異の周辺において、2コピー間で異なる塩基のクラスターが出現した。

この理論的結果の応用として、新規機能獲得遺伝子を検出するアルゴリズムを開発した。そして、それを酵母やハエのゲノムに適用した結果、いくつかの新規機能獲得遺伝子を見つけることが出来た。それらの多くは、数百年前に出現し、その後種分化したほとんど

どの種に共通して存在した。これは、見つかった新規機能獲得遺伝子が進化上非常に重要な役目を担ったということを示唆する (Osada Innan 2008, Takuno Innan 2009)。

最後に、関連の研究の総説を Nature Reviews Genetics 誌にまとめた (Innan Kondrashov 2010)。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 25 件)

- ① Kijima, T. E., and H. Innan, 2010. On the estimation of the insertion time of LTR retrotransposable elements. *Mol. Biol. Evol.* 27: 896-904.
- ② Fawcett, J. A., and H. Innan, 2011. Neutral and non-neutral evolution of duplicated genes with gene conversion. *Genes* 2: 191-209 (belongs to the special issue "Gene Conversion in Duplicated Genes" edited by Innan H.)
- ③ Mansai, S. P., Kado, T., and H. Innan, 2011. The rate and tract length of gene conversion between duplicated genes. *Genes* 2: 313-331. (belongs to the special issue "Gene Conversion in Duplicated Genes" edited by Innan H.)
- ④ Takuno, S., and H. Innan, 2011. Selection fine tunes the expression of microRNA target genes in Arabidopsis. *Mol. Biol. Evol.* (in press)
- ⑤ Sugino, R. P., and H. Innan, 2011. Natural selection on gene order in the genome re-organization process after whole genome duplication of yeast. *Mol. Biol. Evol.* (in press)
- ⑥ Innan, H., and F. Kondrashov, 2010. The evolution of gene duplications: classifying and distinguishing between models. *Nat. Rev. Genet.* 11: 97-108.
- ⑦ Hiwatashi, T., Okabe, H. Tsutsui, T., Hiramatsu, C., Melin, A. D., Oota, H., Schaffner, C. M., Aureli, F., Fedigan, L. M., Innan, H., and S. Kawamura, 2010. An explicit signature of balancing selection for color vision variation in New World Monkeys. *Mol. Biol. Evol.* 27: 453-464.
- ⑧ Mansai, S. P., and H. Innan, 2010. The power of the methods for detecting interlocus gene conversion. *Genetics* 184: 517-527.
- ⑨ Arguello, J. R., Zhang, Y., Kado, T., Fan, C., Zhao, R., Innan, H., Wang, W., Long, M., 2010. Recombination yet inefficient selection along the *Drosophila melanogaster* subgroup's fourth chromosome. *Mol. Biol. Evol.* 27: 848-861.
- ⑩ Takahashi, Y., K. M. Teshima, S. Yokoi, H. Innan and K. Shimamoto, 2009. Variations in Hd1 proteins, *Hd3a* promoters, and *Ehd1* expression levels contribute to diversity of flowering time in cultivated rice. *Proc. Natl. Acad. Sci. USA.* 106: 4555-4560.
- ⑪ Gojobori, J., and H. Innan, 2009, Potential of fish opsin gene duplications to evolve new adaptive functions. *Trends Genet.* 25: 198-202.
- ⑫ Teshima, K. M., and H. Innan, 2009. mbs: modifying Hudson's ms software to generate samples of DNA sequences with a biallelic site under selection. *BMC Bioinformatics* 10: 166.
- ⑬ Innan, H., 2009. Population genetic models of duplicated genes. *Genetica* 137: 19-37.
- ⑭ Nakhleh, L., D. Ruths, and H. Innan, 2009. Gene trees, species trees, and species networks. pp. 275-303. In "Meta-analysis and

Combining Information in Genetics", edited by R. Guerra and D. Allison. Chapman & Hall.

⑮ Takuno, S., and H. Innan, 2009. Selection to maintain paralogous amino acid differences under the pressure of gene conversion in the heat shock protein genes in yeast. *Mol. Biol. Evol.* 26: 2655-2659.

⑯ Takuno, S., and H. Innan, 2008. Evolution of complexity in miRNA mediated gene regulation systems. *Trends in Genet.* 24: 57-59.

⑰ Takahasi, K., and H. Innan, 2008. Inferring the process of human-chimpanzee speciation. *Encyclopedia of Life Sciences*.

⑱ Teshima, KM., and H. Innan, 2008. Neofunctionalization of duplicated genes under the pressure of gene conversion. *Genetics* 178: 1385-1398.

⑲ Gao, L.-z., and H. Innan, 2008. Non-independent domestication of the two rice subspecies, *Oryza sativa* subsp. *indica* and subsp. *japonica*, demonstrated by multilocus microsatellites. *Genetics* 179: 965-976.

⑳ Than, C., R. Sugino, H. Innan, and L. Nakhleh, 2008. Efficient inference of bacterial strain tree from genome-scale multi-locus data. *Bioinformatics* 24: i123-i131.

㉑ Takahasi, K. R., and H. Innan, 2008. The direction of linkage disequilibrium: a new measure based on ancestral-derived status of segregating alleles. *Genetics* 179: 1705-1712.

㉒ Innan, H., and Y. Kim, 2008. Detecting local adaptation using the joint sampling of polymorphism data in the parental and derived populations. *Genetics* 179:

1713-1720.

㉓ Mano, S., and H. Innan, 2008. The Evolutionary Rate of Duplicated Genes under Concerted Evolution. *Genetics* 180: 493-505.

㉔ Takuno, S., T. Nishio, Y. Satta and H. Innan, 2008. Preservation of a pseudogene by gene conversion and diversifying selection. *Genetics* 180: 517-531.

㉕ Osada, N., and H. Innan, 2008. Duplication and gene conversion in the *Drosophila melanogaster* genome. *PLoS Genet.* 4: e1000305.

[学会発表] (計 14 件)

① 印南秀樹 重複遺伝子の共進化 染色体学会 2007/11/27 総合研究大学院大学

② 印南秀樹、高橋亮 連鎖不平衡の方向性 遺伝学会 2007/9/20 名古屋大学

③ 間野修平、印南秀樹 協調進化する多重遺伝子族の進化速度 遺伝学会 2007/9/20 名古屋大学

④ 宅野将平、西尾剛、印南秀樹 重複遺伝子における多様化選択と遺伝子変換の関係 遺伝学会 2007/9/20 名古屋大学

⑤ Innan, H., K. Teshima, N. Osada Signature of neofunctionalization of duplicated genes under the pressure of gene conversion *SMBE* 2008/6/6

⑥ Yamamichi, M., Innan H. No genetic evidence for complex speciation of humans and chimpanzees *SMBE* 2008/6/6

⑦ Innan H. Coevolution of duplicated genes *BGRS* 2008/6/25

⑧ Innan H. Detecting gene conversion between duplicated genes *PAG* 2009 2009/1/10

⑨ 印南秀樹 相互作用のあるときの2遺伝子座モデルとその応用 遺伝学会 2008/9/4 信州大学

⑩ 手島康介、印南秀樹 自然選択の影響を受けた DNA 配列を生成する Coalescent シミュレーションプログラムの開発 遺伝学会 2009/9/17 信州大学

⑪ 五條堀淳、印南秀樹 硬骨魚におけるオプシン遺伝子の重複とその適応的役割

- 遺伝学会 2009/9/17 信州大学
- ⑫ 木島隆之、印南秀樹 LTR レトロトランスポジンの転移時期の推定に対する遺伝子変換の影響 遺伝学会 2009/9/17 信州大学
  - ⑬ 角友之、印南秀樹 交叉と遺伝子変換による組換え率の推定法 遺伝学会 遺伝子変換の検出方法の比較 遺伝学会 2009/9/18 信州大学
  - ⑭ 萬歳明香、印南秀樹 遺伝子変換の検出方法の比較 遺伝学会 2009/9/18 信州大学

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

[その他]

ホームページ等

<http://www.sendou.soken.ac.jp/esb/innan/InnanLab/>

## 6. 研究組織

### (1) 研究代表者

印南 秀樹 (INNAN HIDEKI)

総合研究大学院大学・先導科学研究科・  
准教授

研究者番号：9 0 4 4 4 1 4 0