

Streszczenie rozprawy doktorskiej pt.

**Funkcjonalne retrogeny w genomie człowieka**  
*ang. Functional retrogenes in the human genome*

Joanna Ciomborowska

Promotor: prof. UAM dr hab. Izabela Makałowska

Pracownia Bioinformatyki  
Instytut Biologii Molekularnej i Biotechnologii  
Wydział Biologii  
Uniwersytet im. Adama Mickiewicza w Poznaniu

Poznań 2014

Przedstawiana rozprawa doktorska poświęcona jest funkcjonalnym retrogenom w genomie człowieka, a w szczególności wybranym zagadnieniom związanym z ich ewolucją. Składają się na nią trzy publikacje, w tym jedna praca przeglądowa przedstawiająca dotychczasowy stan wiedzy na temat funkcjonalnych retrogenów. Dwie pozostałe to prace badawcze opisujące nowatorskie analizy i wynikające z nich odkrycia dotyczące powstawania intronów w retrogenach oraz retrogenów, które zastąpiły swoje geny rodzicielskie.

Presented PhD thesis concerns functional retrogenes in the human genome, especially selected aspects of their evolution. It consists of three publications, one review presenting the current state of knowledge in regard to functional retrogenes. Two others are research papers describing innovative analyses and resulting discoveries connected with intron gain in retrogenes and retrogenes that replaced their parental genes.

# **I. Streszczenie**

## 1. Wprowadzenie

Odkrycie retrosekwencji było jednym z najbardziej znaczących wydarzeń w badaniach genomu człowieka i innych organizmów. Retrosekwencje, które początkowo były uznawane za nieistotne z biologicznego punktu widzenia elementy, są obecnie szeroko badane i coraz częściej zauważa się ich znaczenie w kształtowaniu genomów i transkryptomów. Obecnie badania retrosekwencji, w tym retrogenów, obejmują m.in. takie aspekty, jak intensywność zjawiska retropozycji, cechy charakterystyczne retrosekwencji, metody ich identyfikacji, ewolucja, funkcjonalność oraz ekspresja.

Retrogeny to kopie genów powstające w wyniku odwrotnej transkrypcji mRNA i wbudowania powstałego cDNA w sekwencję genomową. Opisany proces nazywany jest retropozycją, a w jego wyniku wieloegzonowy gen rodzicielski daje początek jednoegzonowej retrokopii (Weiner, Deininger et al. 1986). Retrogeny są zwykle nieaktywne i dlatego nazywa się je retroseudogenami lub pseudogenami. Przez wiele lat uznawane były za elementy nieużyteczne, tzw. „śmieciowe” DNA. Od czasu odkrycia pierwszego funkcjonalnego retrogeny w 1985 roku (Soares, Schon et al. 1985) rozpoczęły się badania, które pokazały, że tego typu sekwencje są istotne z punktu widzenia ewolucji genomów.

Pomimo zwiększonego zainteresowania retrogenami, nadal nie wiadomo ile dokładnie jest ich w genomach zwierzęcych, a istniejące dane są zróżnicowane. Przyczyną tych niezgodności są przede wszystkim różnice w metodach ich identyfikacji. Poszukiwanie retrokopii stanowi duże wyzwanie, między innymi ze względu na zróżnicowaną jakość adnotacji poszczególnych genomów, bardzo duże podobieństwo retrosekwencji do genu rodzicielskiego i jego paralogów oraz możliwość zakwalifikowania do retrogenów kopii genów (lub ich fragmentów) powstających w wyniku duplikacji DNA. Uznaje się, że główne cechy retrogenów to brak intronów i elementów regulatorowych, obecność ogona poli(A) oraz powtórzeń otaczających rejon insercji cDNA (Long 2001). W przypadku funkcjonalnych retrogenów dodatkowy wpływ na zróżnicowane wyniki ma samo definiowanie funkcjonalności. Podczas gdy dla jednej z grup badawczych retrogeny były uznawane za funkcjonalne, gdy posiadały co najmniej jedną sekwencję EST (ang. *Expressed Sequence Tag*) lub cDNA jako dowód na ich ekspresję, dla innych ważną była nienaruszona otwarta ramka odczytu (ORF – ang. *Open Reading Frame*) retrogenów (Vinckenbosch, Dupanloup et al. 2006). Często bierze się też pod uwagę tempo zmian pomiędzy retrogenem i genem rodzicielskim określane za pomocą współczynnika Ka/Ks (stosunek liczby mutacji

niesynonimicznych we wszystkich miejscach niesynonimicznych (Ka) do liczby substytucji synonimicznych przypadających na wszystkie miejsca synonimiczne (Ks)) (Betran, Thornton et al. 2002).

Różnice międzygatunkowe kształtują się między innymi poprzez takie zmiany w genomach, jak pojawianie się nowych kopii genów. Ten dodatkowy materiał sprzyja szybkim i intensywnym zmianom ewolucyjnym i dlatego też retropozycję, która jest jednym z głównych źródeł duplikacji genów, uznaje się także jako jeden z najważniejszych mechanizmów sprzyjających powstawaniu zróżnicowania pomiędzy gatunkami. Pomimo, że do tej pory znaleziono stosunkowo niewiele przykładów gatunkowo specyficznych, funkcjonalnych retrogenów, to wiadomo, że ich wpływ na funkcjonowanie organizmu i fenotyp może być znaczący. Bardzo ciekawym przykładem jest retrogen *fgf4* wywołujący chondrodysplazję u psów. Wszystkie psy ras o krótkich nogach są nosicielami tego retrogenu (Parker, VonHoldt et al. 2009). Inny przykład może stanowić myszy retrogen *Rps23rg1*, który odpowiada za regulację poziomu beta-amyloidu oraz fosforylację białka tau, a więc podstawowe zjawiska towarzyszące chorobie Alzheimerera. Odkryto, że mimo intensywnie zachodzącej retropozycji ludzkiego genu rodzicielskiego *RPS23*, *Rps21rg1* występuje tylko u myszy tak jak i drugi funkcjonalny, ulegający ekspresji retrogen *Rps23rg2*. U człowieka natomiast nie zidentyfikowano ortologicznych genów z grupy *Rps23rg*, a wszystkie powstałe retrokopie są najprawdopodobniej pseudogenami (Zhang, Liu et al. 2009). Inne przykłady to specyficzne dla naczelnych funkcjonalne retrogeny, jak *GLUD2* kodujący dehydrogenazę glutaminianową i ulegający ekspresji w mózgu (Burki and Kaessmann 2004) oraz *CDC14Bretro*, którego gen rodzicielski związany jest z cyklem komórkowym (Rosso, Marques et al. 2008).

Na wczesnych etapach badań nad zduplikowanymi genami uważano, że zazwyczaj jeden z nich akumuluje mutacje i staje się niefunkcjonalny (Haldane 1933, Fisher 1935). W związku z tym, naturalną konsekwencją było powszechne traktowanie wszystkich retrokopii, niefunkcjonalnych w momencie powstania, jako pseudogenów. Jednak z czasem okazało się, że „rozluźniona” selekcja i swoboda ewoluowania jakim podlega większość duplikatów, mogą prowadzić nie tylko do pseudogenizacji, ale i nabywania nowych funkcji (Nei 1969). Z czasem opisano dwa nowe zjawiska związane z ewolucją funkcjonalną po duplikacji: neofunkcjonalizacja, w której jedna kopia zdobywa nowe funkcje, a druga zachowuje dotychczasową oraz subfunkcjonalizacja czyli podział już wykształconej funkcji pomiędzy zaangażowane w duplikację geny (Force, Lynch et al. 1999). Jak pokazały nasze

badania, w przypadku retropozycji istnieje także trzecia opcja, czyli zastąpienie genu rodzicielskiego przez retrogen (Ciomborowska, Rosikiewicz et al. 2013).

W wielu dotychczasowych badaniach postulowano, że retrogeny wykazują wąski, często tkankowo-specyficzny zakres działania i że ulegają ekspresji przede wszystkim w jądrach (Vinckenbosch, Dupanloup et al. 2006), (Bai, Casola et al. 2007), (Pan and Zhang 2009) przy jednoczesnej tendencji genów rodzicielskich do ekspresji w wielu tkankach (Marques, Dupanloup et al. 2005), (Bai, Casola et al. 2007), (Potrzebowski, Vinckenbosch et al. 2008). Jest kilka hipotez, które pomagają wyjaśnić tego typu zjawisko. Po pierwsze tłumaczyć to można stanem tzw. hipertranskrypcji w komórkach spermatogenetycznych, w którym odpowiednio zmodyfikowana chromatyna umożliwia transkrypcję tych fragmentów DNA, które w innych warunkach pozostałyby nieaktywne (Marques, Dupanloup et al. 2005), (Chen, Zou et al. 2011). Druga hipoteza wskazuje na możliwość preferencyjnego wbudowywania retrogenów w rejony aktywnej i otwartej chromatyny, szczególnie w pobliżu genów ulegających ekspresji w komórkach zarodkowych. Takie otoczenie wpływa na zwiększoną ekspresję samych retrogenów, w szczególności w jądrach (Fontanillas, Hartl et al. 2007). Jeszcze inny scenariusz wiąże się z teorią tzw. „ucieczki z chromosomu X”. Istnieje dużo przykładów wskazujących na nadreprezentację retrogenów pochodzących od genów rodzicielskich zlokalizowanych na chromosomie X (Betran, Thornton et al. 2002), (Emerson, Kaessmann et al. 2004). Sugeruje się w związku z tym, że „uciekające” po retropozycji na autosomy retrogeny stanowią funkcjonalne odpowiedniki genów źródłowych, które to mogły ulec wyciszeniu w wyniku inaktywacji chromosomu płciowego (Marques, Dupanloup et al. 2005), (Potrzebowski, Vinckenbosch et al. 2008).

Opisane powyżej i inne aspekty badań związanych z retrogenami zostały przeanalizowane i zebrane w rozdziale pod tytułem „Functional retrogenes in animal genomes” w książce *Evolutionary Biology: Mechanisms and Trends*.

## **2. Utrata i powstawanie intronów w retrogenach**

Ogromna większość genów kodujących białka u organizmów eukariotycznych zawiera w swej strukturze introny, których mechanizmy powstawania i wycinania są stosunkowo dobrze poznane i opisane (Chow, Gelinis et al. 1977); (Roy and Gilbert 2006). Wiele dotychczasowych badań wskazało na silny stopień zakonserwowania pozycji intronów w genach nawet daleko spokrewnionych ze sobą organizmów (Rogozin, Wolf et al. 2003); (Carmel, Wolf et al. 2007). Z drugiej strony, metody stosowane w genomice porównawczej

pozwołyły zidentyfikować przykłady utraty, jak i nabywania intronów w toku ewolucji. Zauważono jednak, że powstawanie intronów jest zjawiskiem stosunkowo rzadkim u kręgowców (Loh, Brenner et al. 2007). Co więcej, mimo licznych badań nie zaobserwowano u ssaków żadnego przypadku tzw. „intronizacji”, a więc przekształcenia sekwencji egzonowej w intron (Roy, Fedorov et al. 2003); (Coulombe-Huntington and Majewski 2007). Inne opisywane do tej pory mechanizmy uwzględniały jedynie nabywanie intronu poprzez przyłączenie nowego egzonu (O'Neill, Brennan et al. 1998); (Vinckenbosch, Dupanloup et al. 2006); (Fablet, Bueno et al. 2009).

Nasze bioinformatyczne i eksperymentalne badania retrogenów pozwoliły na znalezienie nowych, gatunkowo specyficznych intronów i pokazały, że powstawanie intronów zachodzi także u ssaków. W wyniku analiz procesów retropozycji genów *RNF113* i *DCAF12* znaleźliśmy dwa jednoegzonowe retrogeny u człowieka *RNF113A* i *DCAF12L1*, a także retrokopie z intronami (*RNF113B* oraz *DCAF12L2*). Weryfikacja struktury wyłonionych retrogenów z intronami nastąpiła między innymi poprzez analizę sekwencji EST. Dzięki temu zidentyfikowaliśmy jeden przypadek tzw. „intronizacji” w specyficznym dla naczelnych retrogenie *RNF113B* oraz dwie niezależne intronizacje w retrogenie *DCAF12L2* – jedna z nich nastąpiła u wspólnego przodka naczelnych i gryzoni, a druga tylko u gryzoni. Dodatkowo, jako pierwsi na świecie, znaleźliśmy i potwierdziliśmy eksperymentalnie retrogeny posiadające warianty splicingowe. Co więcej, prześledziliśmy także profile ekspresji zidentyfikowanych retrogenów w kilkunastu organach i tkankach. Pozwoliło to zaobserwować, że wariant bez intronu ulega ekspresji w wielu tkankach, natomiast ewolucyjnie młodszy wariant z intronem wykazuje tendencję do tkankowo specyficznej ekspresji w jądrach. Ograniczona do jąder ekspresja nowego retrogeny jest zgodna z wcześniejszymi doniesieniami wskazującymi na takie właśnie zachowanie wielu retrokopii (Marques, Dupanloup et al. 2005), jednak my pokazaliśmy, że nie zawsze musi ona być charakterystyczna tylko dla retrogenów pochodzących od genów rodzicielskich zlokalizowanych na chromosomie X (Potrzebowski, Vinckenbosch et al. 2008).

Podsumowując, bardzo ciekawym wnioskiem z naszych badań jest to, że z jednej strony retropozycja powoduje utratę intronów a z drugiej, retrogeny są także miejscem intensywnego powstawania nowych intronów u ssaków. Po raz kolejny potwierdza to, że retrogeny stają się źródłem wielu nowości genomowych. Artykuł opisujący powyższe badania ukazał się w czasopiśmie *Molecular Biology and Evolution* pod tytułem „Primate and rodent specific intron gains and the origin of retrogenes with splice variants”.

### 3. „Osierocone“ retrogeny w genomie człowieka

Przez długi czas retrosekwencje opisywane były jako tzw. „śmieci genomowe“ głównie z tego powodu, że uważano je za nieaktywne i z czasem zanikające geny (pseudogeny) (Mighell, Smith et al. 2000). Dopiero niedawno wykazano, że część retrosekwencji może pozostawać w genomie odgrywając istotną rolę, dając na przykład początek nowym genom (Betran, Wang et al. 2002) lub regulatorowym cząsteczkom RNA (Devor 2006). Aspekty pseudogenizacji i neo- lub subfunkcjonalizacji dominowały w dotychczasowych badaniach retrogenów, ale jak dotąd nigdy nie brano pod uwagę tego, że to nie retrogen, ale jego gen rodzicielski może ulec pseudogenizacji. Teoretycznie takiej sytuacji nie można wykluczyć i dlatego podjęliśmy się trudnego zadania poszukania w genomie człowieka przypadków, w których retrogen całkowicie przejął funkcje swojego genu rodzicielskiego. Wykorzystując nowatorskie podejście do tego zagadnienia oraz szereg narzędzi i danych bioinformatycznych przeprowadziliśmy porównawcze analizy genomów człowieka, kury i niciania, dzięki którym udało się zidentyfikować 25 tzw. „osieroconych” retrogenów w genomie człowieka. Wynik ten jasno pokazał, po raz pierwszy na świecie, że retrogeny nie tylko ulegają pseudogenizacji, neo- lub subfunkcjonalizacji, ale także mogą zastąpić swoje geny rodzicielskie. Prześledzenie historii ewolucyjnej zidentyfikowanych retrogenów pozwoliło z kolei na wykazanie, że większość z nich (14 na 25) powstała i zastąpiła swoje geny rodzicielskie na stosunkowo wczesnych etapach ewolucji zwierząt. Było to zaskakującą informacją, gdyż wcześniejsze doniesienia wskazują na intensywnie zachodzącą retropozycję głównie u ssaków (Moran, Holmes et al. 1996); (Ostlund, Schmitt et al. 2010). Kolejnym ciekawym odkryciem był fakt, że ogromna większość „osieroconych” retrogenów wykazuje szeroki zakres ekspresji w wielu tkankach i organach, co ustalono na podstawie analiz PCR w czasie rzeczywistym. Wspierać to może hipotezę o zastąpieniu genów rodzicielskich, a także stoi w opozycji do wcześniejszych badań wyraźnie sugerujących tkankową specyficzność retrosekwencji, a zwłaszcza silną i często wybiórczą ekspresję w jądrach (Marques, Dupanloup et al. 2005); (Vinckenbosch, Dupanloup et al. 2006); (Potrzebowski, Vinckenbosch et al. 2008). Warto również zaznaczyć, że siedem spośród badanych retrogenów wykazuje powiązania z chorobami u człowieka, takimi jak na przykład rak piersi (Rodriguez, Chen et al. 2007), płasawica Huntingtona (Carnemolla, Fossale et al. 2009) czy cukrzyca typu 2 (Rosengren, Jokubka et al. 2010).



Nasze szczególne zainteresowanie wzbudził retrogen *CHMP1B*, który zastąpił swój gen rodzicielski w genomie człowieka, podczas gdy u myszy występują obydwa w pełni funkcjonalne geny (zarówno wieloegzonowy, jak i jednoegzonowa retrokopia). Co więcej, poprzez analizę potencjalnych miejsc wiązania czynników transkrypcyjnych pokazaliśmy, że ludzki i mysi retrogen mogą być regulowane w podobny sposób, w przeciwieństwie do genu rodzicielskiego u myszy posiadającego zupełnie inny zestaw elementów regulatorowych. Zidentyfikowaliśmy także silniej zachowany wzór miejsc docelowych dla wiązania miRNA w grupie ortologicznych retrogenów *CHMP1B*, niż w przypadku ich istniejących lub zanikających genów rodzicielskich.

Wszystkie te wyniki wskazują na istotną rolę jaką pełnią retrogeny w kształtowaniu cech gatunkowo specyficznych, a funkcjonalne znaczenie retrogenów podkreśla fakt, że ich mutacje mogą prowadzić do rozwoju poważnych chorób. Warto także zwrócić uwagę na to, że zidentyfikowane przez nas „osierocone” retrogeny nie były wcześniej opisane jako geny powstałe w wyniku retropozycji. Wskazuje to na zasadność badań tych dotąd mało poznanych elementów genomów i pokazuje, że analizy retrogenów znacząco przyczyniają się do lepszego poznania genomów i ich ewolucji oraz zrozumienia różnic międzygatunkowych. Opisywane badania zostały opublikowane w czasopiśmie naukowym *Molecular Biology and Evolution*, w artykule pod tytułem „Orphan Retrogenes in the Human Genome”.

#### **4. Podsumowanie**

Głównym celem mojego projektu doktorskiego było przebadanie zjawiska retropozycji u człowieka pod kątem towarzyszących mu zjawisk ewolucyjnych, takich jak nabywanie nowych intronów czy też zastępowanie genów rodzicielskich przez retrogeny. Przeprowadzone analizy pozwoliły na wykrycie tych zjawisk oraz ich scharakteryzowanie. Wszystkie prowadzone przeze mnie dotychczas badania sprzyjają lepszemu poznaniu ludzkiego genomu, który pomimo zakończenia projektu sekwencjonowania i wielu lat analiz wciąż kryje dużo tajemnic. Nasze i inne badania tzw. „śmieciowego DNA“ pokazują, że retrosekwencje mogą być ważne, funkcjonalne, a w wielu procesach wręcz kluczowe. Praca nad tego typu sekwencjami stanowi duże wyzwanie, ale także nieustannie prowokuje do nowych pytań i zadań badawczych. Dlatego też kontynuuję moje badania uczestnicząc w trzech innych projektach związanych z retropozycją. Są to: wielkoskalowa identyfikacja oraz analiza retrogenów i ich genów rodzicielskich w ponad sześćdziesięciu genomach

zwierzęcych; poszukiwanie i charakterystyka kolejnych przykładów intronizacji w retrogenach; a także badanie funkcjonalnych, choć do tej pory uznawanych za pseudogeny, retrogenów specyficznych dla człowieka i istotnych w kształtowaniu różnic międzygatunkowych.

## 5. Bibliografia

- Bai, Y., C. Casola, C. Feschotte and E. Betran (2007). "Comparative genomics reveals a constant rate of origination and convergent acquisition of functional retrogenes in *Drosophila*." *Genome Biol* 8(1): R11.
- Betran, E., K. Thornton and M. Long (2002). "Retroposed new genes out of the X in *Drosophila*." *Genome Res* 12(12): 1854-1859.
- Betran, E., W. Wang, L. Jin and M. Long (2002). "Evolution of the phosphoglycerate mutase processed gene in human and chimpanzee revealing the origin of a new primate gene." *Mol Biol Evol* 19(5): 654-663.
- Burki, F. and H. Kaessmann (2004). "Birth and adaptive evolution of a hominoid gene that supports high neurotransmitter flux." *Nat Genet* 36(10): 1061-1063.
- Carmel, L., Y. I. Wolf, I. B. Rogozin and E. V. Koonin (2007). "Three distinct modes of intron dynamics in the evolution of eukaryotes." *Genome Res* 17(7): 1034-1044.
- Carnemolla, A., E. Fossale, E. Agostoni, S. Michelazzi, R. Calligaris, L. De Maso, G. Del Sal, M. E. MacDonald and F. Persichetti (2009). "Rrs1 is involved in endoplasmic reticulum stress response in Huntington disease." *J Biol Chem* 284(27): 18167-18173.
- Chen, M., M. Zou, B. Fu, X. Li, M. D. Vibranovski, X. Gan, D. Wang, W. Wang, M. Long and S. He (2011). "Evolutionary patterns of RNA-based duplication in non-mammalian chordates." *PLoS One* 6(7): e21466.
- Chow, L. C., R. E. Gelinis, T. R. Broker and R. J. Roberts (1977). "An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA." *Rev Med Virol* 10(6): 362-371; discussion 355-366.
- Ciomborowska, J., W. Rosikiewicz, D. Szklarczyk, W. Makalowski and I. Makalowska (2013). "'Orphan' retrogenes in the human genome." *Mol Biol Evol* 30(2): 384-396.
- Coulombe-Huntington, J. and J. Majewski (2007). "Intron loss and gain in *Drosophila*." *Mol Biol Evol* 24(12): 2842-2850.
- Devor, E. J. (2006). "Primate microRNAs miR-220 and miR-492 lie within processed pseudogenes." *J Hered* 97(2): 186-190.
- Emerson, J. J., H. Kaessmann, E. Betran and M. Long (2004). "Extensive gene traffic on the mammalian X chromosome." *Science* 303(5657): 537-540.
- Fablet, M., M. Bueno, L. Potrzebowski and H. Kaessmann (2009). "Evolutionary origin and functions of retrogene introns." *Mol Biol Evol* 26(9): 2147-2156.
- Fisher R (1935) The sheltering of lethals. *AmNat* 69:446-455.

Fontanillas, P., D. L. Hartl and M. Reuter (2007). "Genome organization and gene expression shape the transposable element distribution in the *Drosophila melanogaster* euchromatin." *PLoS Genet* 3(11): e210.

Force, A., M. Lynch, F. B. Pickett, A. Amores, Y. L. Yan and J. Postlethwait (1999). "Preservation of duplicate genes by complementary, degenerative mutations." *Genetics* 151(4): 1531-1545.

Haldane J (1933) The part played by recurrent mutation in evolution. *Am Nat* 67:5-19.

Loh, Y. H., S. Brenner and B. Venkatesh (2007). "Investigation of loss and gain of introns in the compact genomes of pufferfishes (*Fugu* and *Tetraodon*)." *Mol Biol Evol* 25(3): 526-535.

Long, M. (2001). "Evolution of novel genes." *Curr Opin Genet Dev* 11(6): 673-680.

Marques, A. C., I. Dupanloup, N. Vinckenbosch, A. Reymond and H. Kaessmann (2005). "Emergence of young human genes after a burst of retroposition in primates." *PLoS Biol* 3(11): e357.

Mighell, A. J., N. R. Smith, P. A. Robinson and A. F. Markham (2000). "Vertebrate pseudogenes." *FEBS Lett* 468(2-3): 109-114.

Moran, J. V., S. E. Holmes, T. P. Naas, R. J. DeBerardinis, J. D. Boeke and H. H. Kazazian, Jr. (1996). "High frequency retrotransposition in cultured mammalian cells." *Cell* 87(5): 917-927.

Nei, M. (1969). "Gene duplication and nucleotide substitution in evolution." *Nature* 221(5175): 40-42.

O'Neill, R. J., F. E. Brennan, M. L. Delbridge, R. H. Crozier and J. A. Graves (1998). "De novo insertion of an intron into the mammalian sex determining gene, SRY." *Proc Natl Acad Sci U S A* 95(4): 1653-1657.

Ostlund, G., T. Schmitt, K. Forslund, T. Kostler, D. N. Messina, S. Roopra, O. Frings and E. L. Sonnhammer (2010). "InParanoid 7: new algorithms and tools for eukaryotic orthology analysis." *Nucleic Acids Res* 38(Database issue): D196-203.

Pan, D. and L. Zhang (2009). "Burst of young retrogenes and independent retrogene formation in mammals." *PLoS One* 4(3): e5040.

Parker, H. G., B. M. VonHoldt, P. Quignon, E. H. Margulies, S. Shao, D. S. Mosher, T. C. Spady, A. Elkahoun, M. Cargill, P. G. Jones, C. L. Maslen, G. M. Acland, N. B. Sutter, K. Kuroki, C. D. Bustamante, R. K. Wayne and E. A. Ostrander (2009). "An expressed *fgf4* retrogene is associated with breed-defining chondrodysplasia in domestic dogs." *Science* 325(5943): 995-998.

Potrzebowski, L., N. Vinckenbosch, A. C. Marques, F. Chalmel, B. Jegou and H. Kaessmann (2008). "Chromosomal gene movements reflect the recent origin and biology of therian sex chromosomes." *PLoS Biol* 6(4): e80.

Rodriguez, V., Y. Chen, A. Elkahloun, A. Dutra, E. Pak and S. Chandrasekharappa (2007). "Chromosome 8 BAC array comparative genomic hybridization and expression analysis identify amplification and overexpression of TRMT12 in breast cancer." *Genes Chromosomes Cancer* 46(7): 694-707.

Rogozin, I. B., Y. I. Wolf, A. V. Sorokin, B. G. Mirkin and E. V. Koonin (2003). "Remarkable interkingdom conservation of intron positions and massive, lineage-specific intron loss and gain in eukaryotic evolution." *Curr Biol* 13(17): 1512-1517.

Rosengren, A. H., R. Jokubka, D. Tojjar, C. Granhall, O. Hansson, D. Q. Li, V. Nagaraj, T. M. Reinbothe, J. Tuncel, L. Eliasson, L. Groop, P. Rorsman, A. Salehi, V. Lyssenko, H. Luthman and E. Renstrom (2010). "Overexpression of alpha2A-adrenergic receptors contributes to type 2 diabetes." *Science* 327(5962): 217-220.

Rosso, L., A. C. Marques, M. Weier, N. Lambert, M. A. Lambot, P. Vanderhaeghen and H. Kaessmann (2008). "Birth and rapid subcellular adaptation of a hominoid-specific CDC14 protein." *PLoS Biol* 6(6): e140.

Roy, S. W., A. Fedorov and W. Gilbert (2003). "Large-scale comparison of intron positions in mammalian genes shows intron loss but no gain." *Proc Natl Acad Sci U S A* 100(12): 7158-7162.

Roy, S. W. and W. Gilbert (2006). "The evolution of spliceosomal introns: patterns, puzzles and progress." *Nat Rev Genet* 7(3): 211-221.

Soares, M. B., E. Schon, A. Henderson, S. K. Karathanasis, R. Cate, S. Zeitlin, J. Chirgwin and A. Efstratiadis (1985). "RNA-mediated gene duplication: the rat preproinsulin I gene is a functional retroposon." *Mol Cell Biol* 5(8): 2090-2103.

Vinckenbosch, N., I. Dupanloup and H. Kaessmann (2006). "Evolutionary fate of retroposed gene copies in the human genome." *Proc Natl Acad Sci U S A* 103(9): 3220-3225.

Weiner, A. M., P. L. Deininger and A. Efstratiadis (1986). "Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information." *Annu Rev Biochem* 55: 631-661.

Zhang, Y. W., S. Liu, X. Zhang, W. B. Li, Y. Chen, X. Huang, L. Sun, W. Luo, W. J. Netzer, R. Threadgill, G. Wiegand, R. Wang, S. N. Cohen, P. Greengard, F. F. Liao, L. Li and H. Xu (2009). "A functional mouse retroposed gene Rps23r1 reduces Alzheimer's beta-amyloid levels and tau phosphorylation." *Neuron* 64(3): 328-340.

## **II. Summary**

## 1. Introduction

The discovery of retrosequences was one of the most significant events in genome analyses. Retrosequences, previously described as completely useless and biologically unimportant elements, now are widely analyzed and their roles in shaping animal genomes and transcriptomes become more and more visible. Currently retrosequences are studied in many aspects, including frequency of retroposition, characteristics of retrosequences, identification methods, their evolution, functionality and expression.

Retrogenes are copies of genes originated from reverse transcription of mRNA and incorporation of cDNA into a genomic sequence. This process is called retroposition and it results in a formation of a single-exon copy from a multi-exon parental gene (Weiner, Deininger et al. 1986). Retrogenes are usually inactive and therefore are commonly called retropseudogenes or just pseudogenes. For many years they have been considered as useless, so called “junk DNA”. Nevertheless, since the discovery of the first functional retrogene in 1985 (Soares, Schon et al. 1985) the interest in retrogenes increased and subsequent studies revealed that this type of sequences is important from the evolutionary point of view.

Despite the increased interest in retrogenes, it is still not known how many of them are there in animal genomes and existing data are diversified. This variance may be explained mostly by different methods of retrogenes identification. Searching for retrocopies is a big challenge, mainly because of diverse quality of genomic annotations and high similarity between retrosequences, parental genes and their paralogs. Moreover, there is a strong possibility that genes (or their fragments) duplicated via DNA-based mechanisms can be wrongly qualified as retrogenes. The most distinctive attributes of retrogenes are lack of introns and regulatory elements, presence of poly(A) tail, and repeats located near to the insertion region (Long 2001). In case of functional retrogenes, there is an additional important aspect affecting results, which is the definition of functionality. Whereas for one research group, retrogenes can be considered as functional when their expression can be proved by at least one EST (*Expressed Sequence Tag*) or cDNA sequence, for the others the existence of intact ORF (*Open Reading Frame*) (Vinckenbosch, Dupanloup et al. 2006) or small rate of the evolutionary changes between retrogene and parental gene measured by Ka/Ks ratio (the ratio of non-synonymous substitutions (Ka) to synonymous substitutions (Ks)) (Betran, Thornton et al. 2002) were the most important factors.

Interspecies differences are shaped by various changes in genomes like gain of new gene copies for example. This additional genetic material enhances evolutionary changes and therefore retroposition, as crucial source of duplicated genes, is considered to be one of the most essential processes responsible for interspecies diversification. Although relatively few examples of species-specific functional retrogenes were reported, it is known that their impact on the function of the organism and its phenotype can be significant. A very interesting example is the retrogene *fgf4* responsible for the dogs' chondrodysplasia. All breeds with short legs are carriers of this retrogene (Parker, VonHoldt et al. 2009). Another case is the *Rps23rg1* retrogene in mouse that is responsible for the regulation of beta-amyloid level and tau protein phosphorylation, basic phenomena related to Alzheimer's disease. It was discovered that in spite of the intense retroposition of human parental gene (*RPS23*), *Rps23rg1* and second functional and expressed retrogene *Rps23rg2* exist only in mouse. In human there are no orthologous genes from *Rps23rg* family and all retrocopies present in the genome are most probably pseudogenes (Zhang, Liu et al. 2009). Other examples of primate-specific functional retrogenes include *GLUD2* coding glutamate dehydrogenase and expressed in brain (Burki and Kaessmann 2004) and *CDC14Bretro*, which originated from parental gene related to cell cycle (Rosso, Marques et al. 2008).

In the early studies of duplicated genes evolution it was postulated that usually one of the duplicates accumulates mutations and becomes nonfunctional (Haldane 1933); (Fisher 1935). Consequently, all retrocopies, which are nonfunctional at the moment of their origin, were considered as pseudogenes. However, it occurred that "relaxed" selection and evolutionary freedom which are characteristic for majority of duplicates, may lead not only to pseudogenization but also to acquisition of new functions (Nei 1969). Over the time two new phenomena related to functional evolution after duplication were described: neofunctionalization, where one copy acquires a new function and the other one keeps the original one and subfunctionalization when maintained function is shared between duplicated genes (Force, Lynch et al. 1999). As our studies showed, there is also another possibility: the retrogene may replace its parent (Ciomborowska, Rosikiewicz et al. 2013).

In many recent studies it has been suggested that retrogenes tend to exhibit narrow, often tissue-specific expression pattern and are expressed mainly in testes (Vinckenbosch, Dupanloup et al. 2006); (Bai, Casola et al. 2007); (Pan and Zhang 2009). At the same time, the general tendency for parental genes to be broadly expressed was shown (Marques, Dupanloup et al. 2005); (Bai, Casola et al. 2007); (Potrzebowski, Vinckenbosch et al. 2008). There are a few possible hypotheses interpreting this phenomenon. Firstly, it can be explained



by so called hypertranscription state in spermatogenic cells, in which modified chromatin enables transcription of DNA that usually remains inactive (Marques, Dupanloup et al. 2005), (Chen, Zou et al. 2011). The second hypothesis is based on the idea of preferential insertion of retrogenes into active and open chromatin, especially near germline-expressed genes. Such surroundings have a big impact on higher expression level of retrogenes, particularly in testes (Fontanillas, Hartl et al. 2007). Another scenario involves so called “out-of-X chromosome escape” theory. There are plenty of examples showing the overrepresentation of retrogenes originated from parental genes located on chromosome X (Betran, Thornton et al. 2002), (Emerson, Kaessmann et al. 2004). It was suggested that retrocopies escaping from chromosome X may work as autosomal counterparts of their source genes, which could be silenced during male meiotic sex chromosome inactivation (Marques, Dupanloup et al. 2005), (Potrzebowski, Vinckenbosch et al. 2008).

These above mentioned, together with other aspects related to retrogenes, were analyzed and described in the chapter entitled “Functional retrogenes in animal genomes” and published in the book *Evolutionary Biology: Mechanisms and Trends*.

## **2. Gain and loss of introns in retrogenes**

Vast majority of protein-coding genes in eukaryotes contains introns, whose origin and splicing machinery are relatively well known and described (Chow, Gelinas et al. 1977); (Roy and Gilbert 2006). Many studies have revealed a strong level of intron position conservation even in distant organisms (Rogozin, Wolf et al. 2003); (Carmel, Wolf et al. 2007). On the other hand, comparative genomics methods allowed to identify many examples of intron loss and gain during the evolution but it was found that intron gain is a very rare event in vertebrates (Loh, Brenner et al. 2007). Moreover, no cases of so called “intronization” - a transformation of exonic sequence into intron - have been discovered in mammals (Roy, Fedorov et al. 2003); (Coulombe-Huntington and Majewski 2007). The only described mechanisms of intron acquisition were associated with new exon capture (O'Neill, Brennan et al. 1998); (Vinckenbosch, Dupanloup et al. 2006); (Fablet, Bueno et al. 2009).

Our bioinformatics and experimental analyses of retrogenes allowed identification of novel, species-specific introns and showed that intron formation process is active also in mammals. As a result of studies of genes *RNF113* and *DCAF12* retroposition, we identified in the human genome two single-exon retrogenes (*RNF113A*, *DCAF12L1*) as well as retrocopies with introns (*RNF113B* and *DCAF12L2*). Structural verification of obtained

candidates was performed through the analysis of EST sequences. As an outcome we identified one case of “intronization” in primate-specific retrogene *RNF113B* and two independent “intronization” events in the retrogene *DCAF12L2* – one took place in the common ancestor of primates and rodents and another one in the rodent lineage. Additionally, as a first group in the world, we found and experimentally confirmed retrogenes with splicing variants. What is more, we examined expression profiles of those retrogenes in over a dozen of organs and tissues and revealed that a variant without an intron is widely expressed, while a new splicing form containing the intron shows a tendency for tissue-specific expression in testes. Limited to testes expression of the recently originated retrogene variant confirms earlier reports describing such expression pattern of many retrocopies (Marques, Dupanloup et al. 2005). Nevertheless, we demonstrated that not only retrogenes originated from chromosome X show such tendency (Potrzebowski, Vinckenbosch et al. 2008).

Summing up, one of the most interesting conclusions coming from our research is that although retroposition causes intron loss, retrogenes might be regarded as a place of intense intron gain in mammals. This confirms that retrogenes can be viewed as a source of genomic novelties. The article with all described results, entitled “Primate and rodent specific intron gains and the origin of retrogenes with splice variants” was published in the journal *Molecular Biology and Evolution*.

### **3. „Orphan” retrogenes in the human genome**

Retrosequences for a long time have been considered as “genomic junk” mainly because they were regarded as inactive and disappearing over the time genes (pseudogenes) (Mighell, Smith et al. 2000). Only recently it has been showed that some retrosequences may remain in the genome and play a crucial role by giving birth to new genes (Betran, Wang et al. 2002) or regulatory RNAs (Devor 2006). Aspects of pseudogenization together with neo- and sub-functionalization were dominating in all previous analyses and so far nobody took into consideration a situation in which not the retrogene but the parental gene is pseudogenized. Theoretically we cannot exclude such scenario and therefore we decided to take on this challenging task, focusing on identifying in the human genome cases where retrogene overtook the function of its parent. Utilizing innovative approach and a number of bioinformatics tools we performed comparative analyses of human, chicken and nematode genomes, which led us to identification of 25 “orphan” retrogenes in the human genome. These results clearly showed, for the first time, that retrogenes not only can be

pseudogenized, neo- or subfunctionalized, but also are able to replace their progenitors. Analysis of the evolutionary history of identified retrogenes showed that majority of them (14 out of 25) originated and replaced their parents in the early stages of animal evolution. It was a surprising result because previous reports suggested very intense retroposition mainly in mammals (Moran, Holmes et al. 1996); (Ostlund, Schmitt et al. 2010). Another fascinating discovery showed that vast majority of “orphan” retrogenes have a broad range of expression, which we demonstrated with real-time PCR experiments. It may support our hypothesis about parental gene replacement and stays in contrary to earlier research suggesting tissue-specificity of retrogenes (specially strong and specific expression in testes) (Marques, Dupanloup et al. 2005); (Vinckenbosch, Dupanloup et al. 2006); (Potrzebowski, Vinckenbosch et al. 2008). It is worth emphasizing that seven of the identified retrogenes are related to human diseases, such as breast cancer (Rodriguez, Chen et al. 2007), Huntington’s disease (Carnemolla, Fossale et al. 2009), type 2 diabetes (Rosengren, Jokubka et al. 2010).

A case of retrogene *CHMP1B* aroused our special interest. This gene replaced its source gene in the human genome, while in mouse two fully functional genes (both multi-exon and single-exon copy) exist. Detailed analyses of potential transcription factor binding sites showed that human and mouse retrogenes may be regulated in a similar way, in contrast to mouse parental gene having completely different set of regulatory elements. We also identified stronger conservation of target binding sites for miRNA among orthologous retrogenes *CHMP1B*, than in case of their existing or disappearing parental genes.

All these results demonstrate a crucial role of retrogenes in shaping species-specific traits. The fact that their mutations are often leading to diseases underlines their functional importance. It is also worth noticing that majority of identified by us retrogenes haven’t been earlier described as sequences created via retroposition. This strongly highlights the necessity of further analysis of these still little known genomic elements and shows that retrogene analysis can greatly enrich our knowledge about genomes and their evolution as well as improve the understanding of interspecies differences. Described above results were published in the article “*Orphan Retrogenes in the Human Genome*” in the journal *Molecular Biology and Evolution*.

#### **4. Summary**

The main goal of my PhD project was the analysis of retroposition in the human genome with a special focus on accompanying evolutionary phenomena such as intron gain and replacement of parental genes by retrogenes. All performed research led not only to the discovery of abovementioned processes but also to our better understanding of the human genome, which despite the fact of finished sequencing and great number of analyses still hides many secrets. Our and other studies of so called “junk DNA” show that retrosequences might be important, functional and crucial in various processes. Working with such sequences is a big challenge but, at the same time, it constantly gives rise to new questions and research tasks. Therefore, I continue my studies on retrogenes by taking a part in three other projects related to retroposition. These are: identification and large-scale analyses of retrogenes and their progenitors in over sixty animal genomes; searching for next “intronization” events in retrogenes; analysis of human-specific functional retrogenes, considered so far as pseudogenes, which could be involved in formation of interspecies differences.

## 5. Bibliography

- Bai, Y., C. Casola, C. Feschotte and E. Betran (2007). "Comparative genomics reveals a constant rate of origination and convergent acquisition of functional retrogenes in *Drosophila*." *Genome Biol* 8(1): R11.
- Betran, E., K. Thornton and M. Long (2002). "Retroposed new genes out of the X in *Drosophila*." *Genome Res* 12(12): 1854-1859.
- Betran, E., W. Wang, L. Jin and M. Long (2002). "Evolution of the phosphoglycerate mutase processed gene in human and chimpanzee revealing the origin of a new primate gene." *Mol Biol Evol* 19(5): 654-663.
- Burki, F. and H. Kaessmann (2004). "Birth and adaptive evolution of a hominoid gene that supports high neurotransmitter flux." *Nat Genet* 36(10): 1061-1063.
- Carmel, L., Y. I. Wolf, I. B. Rogozin and E. V. Koonin (2007). "Three distinct modes of intron dynamics in the evolution of eukaryotes." *Genome Res* 17(7): 1034-1044.
- Carnemolla, A., E. Fossale, E. Agostoni, S. Michelazzi, R. Calligaris, L. De Maso, G. Del Sal, M. E. MacDonald and F. Persichetti (2009). "Rrs1 is involved in endoplasmic reticulum stress response in Huntington disease." *J Biol Chem* 284(27): 18167-18173.
- Chen, M., M. Zou, B. Fu, X. Li, M. D. Vibranovski, X. Gan, D. Wang, W. Wang, M. Long and S. He (2011). "Evolutionary patterns of RNA-based duplication in non-mammalian chordates." *PLoS One* 6(7): e21466.
- Chow, L. C., R. E. Gelinas, T. R. Broker and R. J. Roberts (1977). "An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA." *Rev Med Virol* 10(6): 362-371; discussion 355-366.
- Ciomborowska, J., W. Rosikiewicz, D. Szklarczyk, W. Makalowski and I. Makalowska (2013). "'Orphan' retrogenes in the human genome." *Mol Biol Evol* 30(2): 384-396.
- Coulombe-Huntington, J. and J. Majewski (2007). "Intron loss and gain in *Drosophila*." *Mol Biol Evol* 24(12): 2842-2850.
- Devor, E. J. (2006). "Primate microRNAs miR-220 and miR-492 lie within processed pseudogenes." *J Hered* 97(2): 186-190.
- Emerson, J. J., H. Kaessmann, E. Betran and M. Long (2004). "Extensive gene traffic on the mammalian X chromosome." *Science* 303(5657): 537-540.
- Fablet, M., M. Bueno, L. Potrzebowski and H. Kaessmann (2009). "Evolutionary origin and functions of retrogene introns." *Mol Biol Evol* 26(9): 2147-2156.
- Fisher R (1935) The sheltering of lethals. *AmNat* 69:446-455.

Fontanillas, P., D. L. Hartl and M. Reuter (2007). "Genome organization and gene expression shape the transposable element distribution in the *Drosophila melanogaster* euchromatin." *PLoS Genet* 3(11): e210.

Force, A., M. Lynch, F. B. Pickett, A. Amores, Y. L. Yan and J. Postlethwait (1999). "Preservation of duplicate genes by complementary, degenerative mutations." *Genetics* 151(4): 1531-1545.

Haldane J (1933) The part played by recurrent mutation in evolution. *Am Nat* 67:5-19.

Loh, Y. H., S. Brenner and B. Venkatesh (2007). "Investigation of loss and gain of introns in the compact genomes of pufferfishes (*Fugu* and *Tetraodon*)." *Mol Biol Evol* 25(3): 526-535.

Long, M. (2001). "Evolution of novel genes." *Curr Opin Genet Dev* 11(6): 673-680.

Marques, A. C., I. Dupanloup, N. Vinckenbosch, A. Reymond and H. Kaessmann (2005). "Emergence of young human genes after a burst of retroposition in primates." *PLoS Biol* 3(11): e357.

Mighell, A. J., N. R. Smith, P. A. Robinson and A. F. Markham (2000). "Vertebrate pseudogenes." *FEBS Lett* 468(2-3): 109-114.

Moran, J. V., S. E. Holmes, T. P. Naas, R. J. DeBerardinis, J. D. Boeke and H. H. Kazazian, Jr. (1996). "High frequency retrotransposition in cultured mammalian cells." *Cell* 87(5): 917-927.

Nei, M. (1969). "Gene duplication and nucleotide substitution in evolution." *Nature* 221(5175): 40-42.

O'Neill, R. J., F. E. Brennan, M. L. Delbridge, R. H. Crozier and J. A. Graves (1998). "De novo insertion of an intron into the mammalian sex determining gene, *SRY*." *Proc Natl Acad Sci U S A* 95(4): 1653-1657.

Ostlund, G., T. Schmitt, K. Forslund, T. Kostler, D. N. Messina, S. Roopra, O. Frings and E. L. Sonnhammer (2010). "InParanoid 7: new algorithms and tools for eukaryotic orthology analysis." *Nucleic Acids Res* 38(Database issue): D196-203.

Pan, D. and L. Zhang (2009). "Burst of young retrogenes and independent retrogene formation in mammals." *PLoS One* 4(3): e5040.

Parker, H. G., B. M. VonHoldt, P. Quignon, E. H. Margulies, S. Shao, D. S. Mosher, T. C. Spady, A. Elkahoun, M. Cargill, P. G. Jones, C. L. Maslen, G. M. Acland, N. B. Sutter, K. Kuroki, C. D. Bustamante, R. K. Wayne and E. A. Ostrander (2009). "An expressed *fgf4* retrogene is associated with breed-defining chondrodysplasia in domestic dogs." *Science* 325(5943): 995-998.

Potrzebowski, L., N. Vinckenbosch, A. C. Marques, F. Chalmel, B. Jegou and H. Kaessmann (2008). "Chromosomal gene movements reflect the recent origin and biology of therian sex chromosomes." *PLoS Biol* 6(4): e80.

- Rodriguez, V., Y. Chen, A. Elkahlon, A. Dutra, E. Pak and S. Chandrasekharappa (2007). "Chromosome 8 BAC array comparative genomic hybridization and expression analysis identify amplification and overexpression of TRMT12 in breast cancer." *Genes Chromosomes Cancer* 46(7): 694-707.
- Rogozin, I. B., Y. I. Wolf, A. V. Sorokin, B. G. Mirkin and E. V. Koonin (2003). "Remarkable interkingdom conservation of intron positions and massive, lineage-specific intron loss and gain in eukaryotic evolution." *Curr Biol* 13(17): 1512-1517.
- Rosengren, A. H., R. Jokubka, D. Tojjar, C. Granhall, O. Hansson, D. Q. Li, V. Nagaraj, T. M. Reinbothe, J. Tuncel, L. Eliasson, L. Groop, P. Rorsman, A. Salehi, V. Lyssenko, H. Luthman and E. Renstrom (2010). "Overexpression of alpha2A-adrenergic receptors contributes to type 2 diabetes." *Science* 327(5962): 217-220.
- Rosso, L., A. C. Marques, M. Weier, N. Lambert, M. A. Lambot, P. Vanderhaeghen and H. Kaessmann (2008). "Birth and rapid subcellular adaptation of a hominoid-specific CDC14 protein." *PLoS Biol* 6(6): e140.
- Roy, S. W., A. Fedorov and W. Gilbert (2003). "Large-scale comparison of intron positions in mammalian genes shows intron loss but no gain." *Proc Natl Acad Sci U S A* 100(12): 7158-7162.
- Roy, S. W. and W. Gilbert (2006). "The evolution of spliceosomal introns: patterns, puzzles and progress." *Nat Rev Genet* 7(3): 211-221.
- Soares, M. B., E. Schon, A. Henderson, S. K. Karathanasis, R. Cate, S. Zeitlin, J. Chirgwin and A. Efstratiadis (1985). "RNA-mediated gene duplication: the rat preproinsulin I gene is a functional retroposon." *Mol Cell Biol* 5(8): 2090-2103.
- Vinckenbosch, N., I. Dupanloup and H. Kaessmann (2006). "Evolutionary fate of retroposed gene copies in the human genome." *Proc Natl Acad Sci U S A* 103(9): 3220-3225.
- Weiner, A. M., P. L. Deininger and A. Efstratiadis (1986). "Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information." *Annu Rev Biochem* 55: 631-661.
- Zhang, Y. W., S. Liu, X. Zhang, W. B. Li, Y. Chen, X. Huang, L. Sun, W. Luo, W. J. Netzer, R. Threadgill, G. Wiegand, R. Wang, S. N. Cohen, P. Greengard, F. F. Liao, L. Li and H. Xu (2009). "A functional mouse retroposed gene Rps23r1 reduces Alzheimer's beta-amyloid levels and tau phosphorylation." *Neuron* 64(3): 328-340.