

Using Evolutionary Principles to Discriminate the Non-Redundant Functions of Related Proteins Critical to Innate Immunity and Cancer Biology

Caitlyn Vlasschaert

Work for M.Sc. in Biochemistry and Bioinformatics

Supervisors: Dr. Douglas Gray and Dr. Xuhua Xia
Ottawa Hospital Research Institute
University of Ottawa

Conflicts of Interest

- I have no conflicts to declare.

Using Evolutionary Principles to Discriminate the **Non-Redundant Functions** of Related Proteins Critical to **Innate Immunity and Cancer Biology**

Related proteins: Gene duplication and alternative splicing

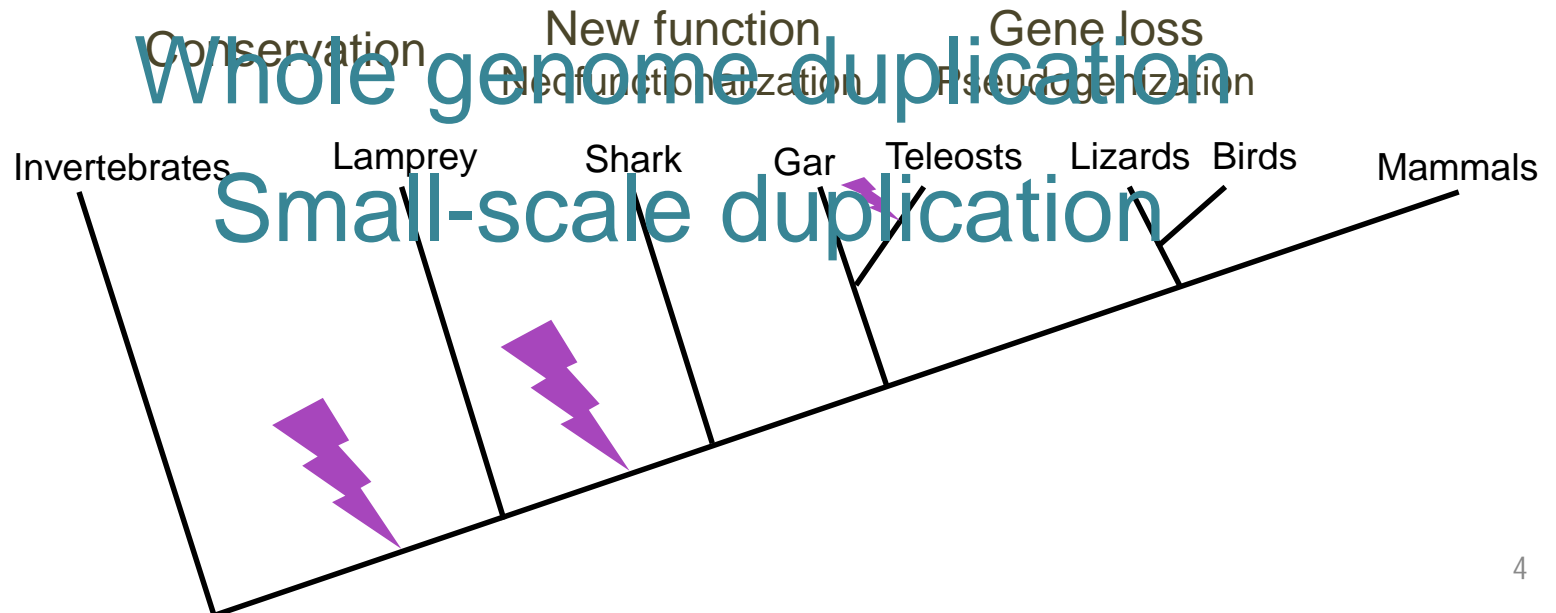
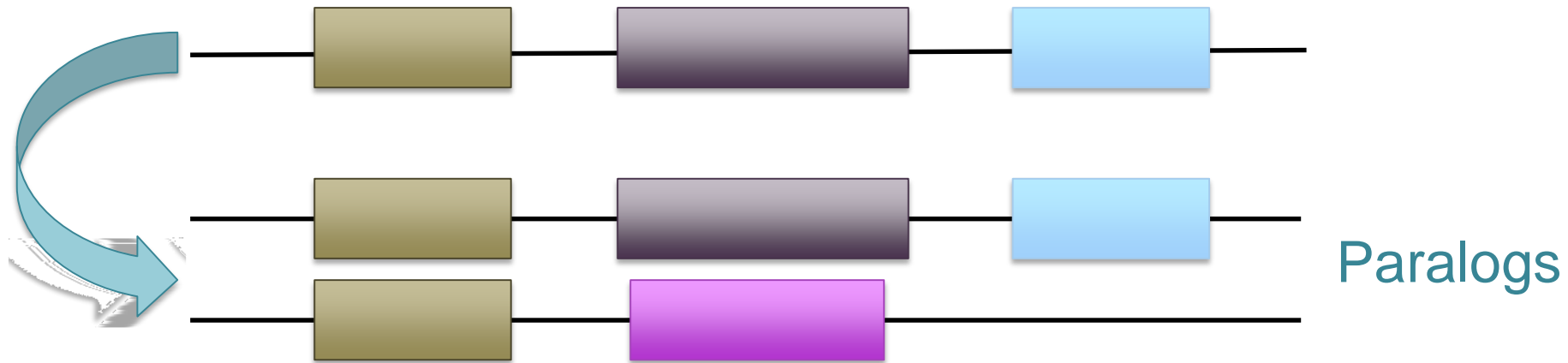
Deubiquitinating enzymes – USP4, USP15, USP11

Evolutionary principles: Computational tools to infer how natural selection shapes proteins over time as different species diverge from the tree of life

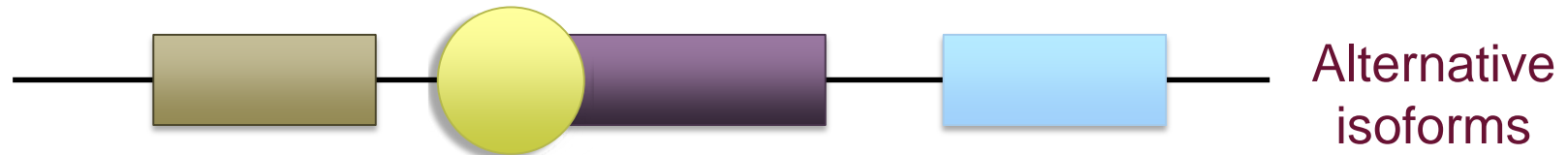
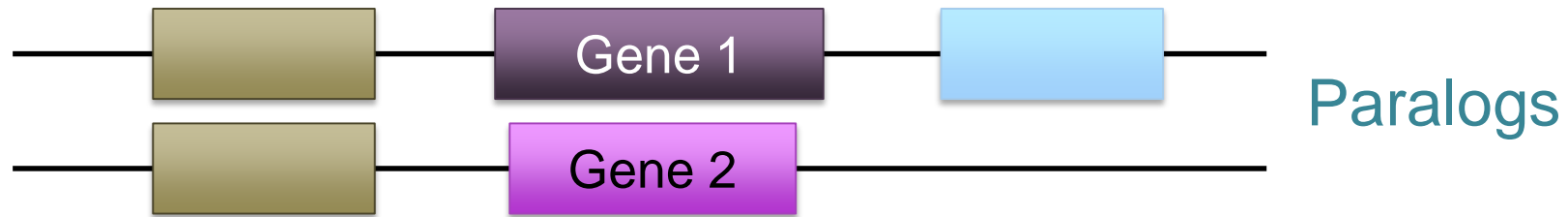


Guiding basic science research

Gene duplication



Alternative splicing



pre-mRNA:

AUGACAAUUCCC**GUAUCGAAUCACCAUUGACAC**ACAUGACCUGUACAU...

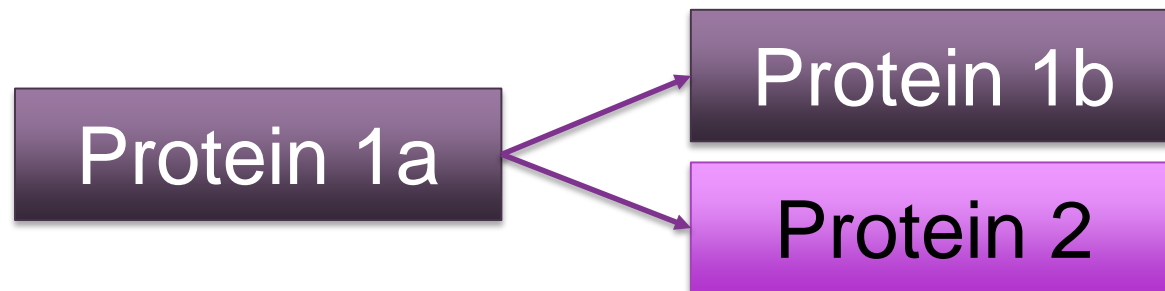
1a

mature mRNA:

AUGACAAUUCCCACAUACAUGACCUGUACAU...

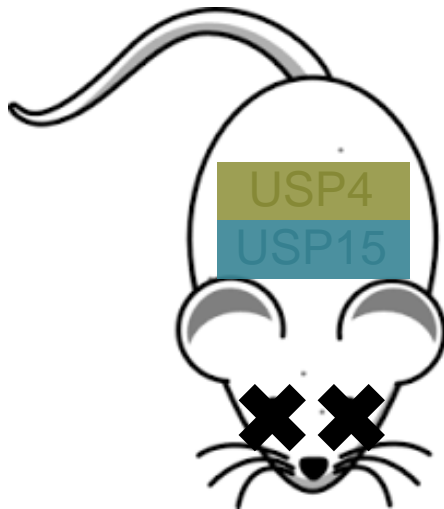
1b

AUGACAAUUCCC**GUAUCGAAUCACCAUUGACAC**ACAUGACCUGUACAU...

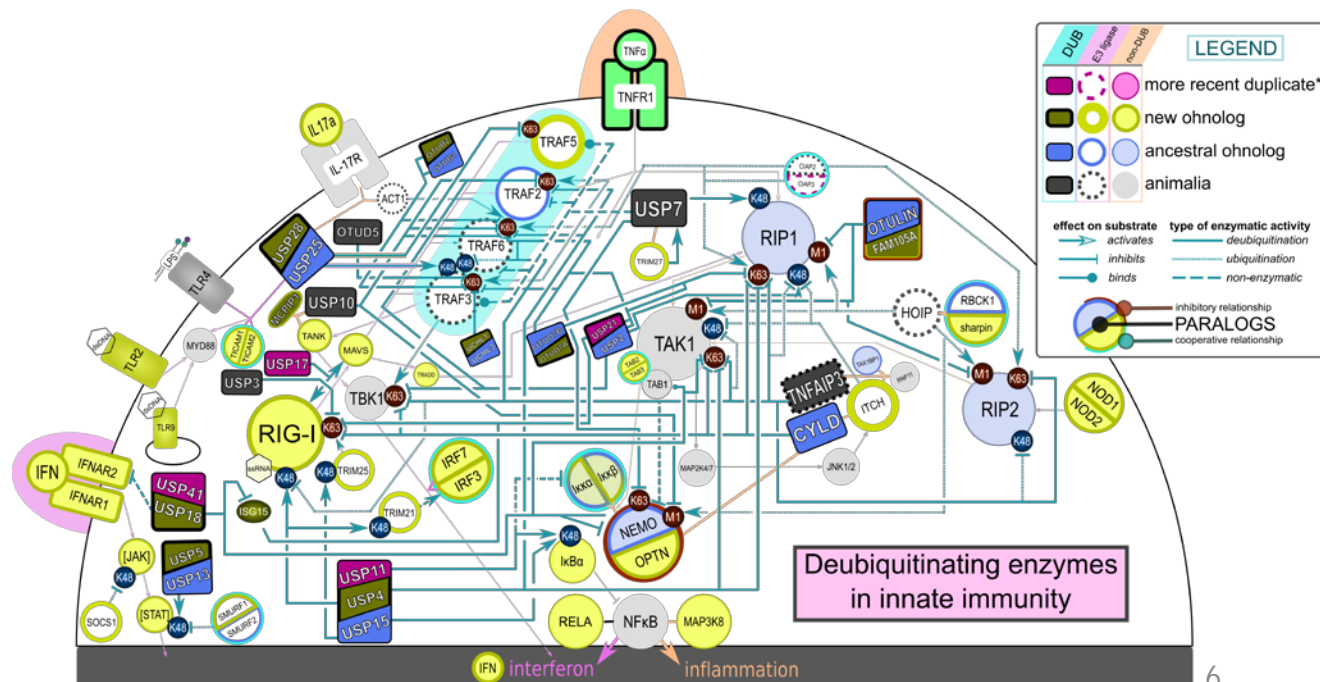


Using Evolutionary Principles to Discriminate the Non-Redundant Functions of Related Proteins Critical to Innate Immunity and Cancer Biology

Paralogs and alternative isoforms often retain overlapping roles. What (if anything) makes them unique and essential?

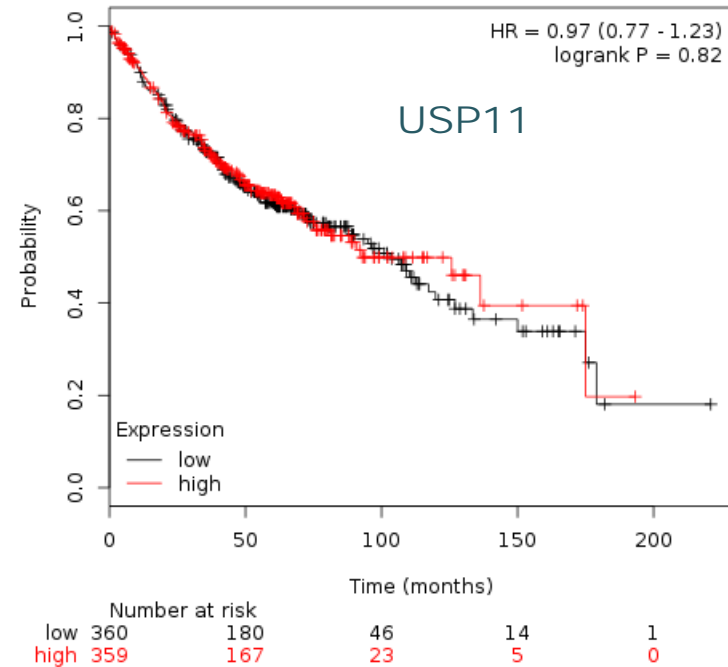
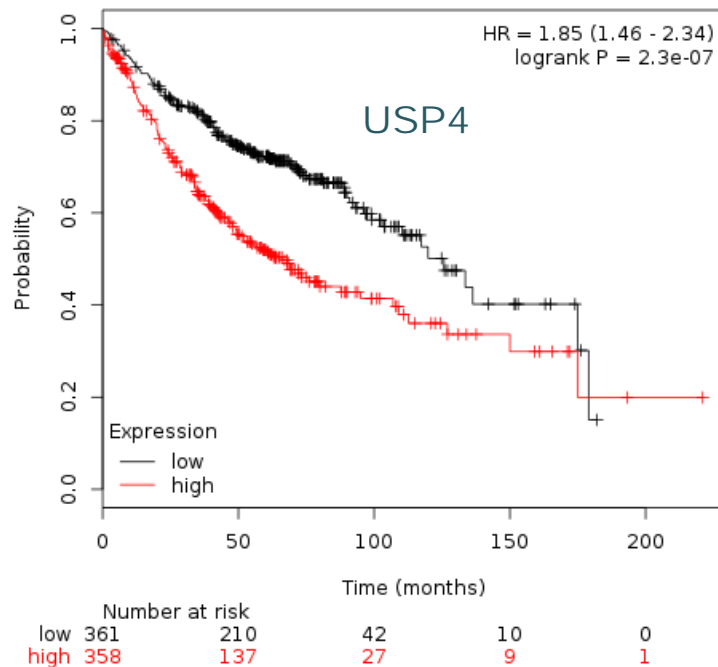


	WT	HET	Null
WT	26	31	17
HET	23	42	12
Null	9	25	0
Total: 185			

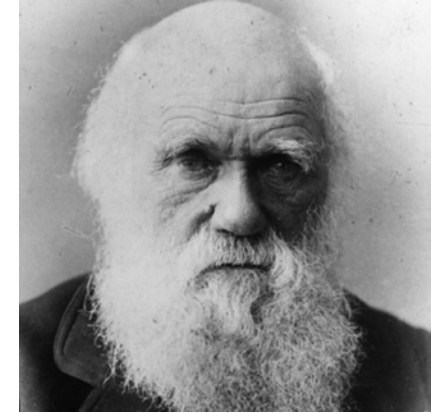


The opposing correlations of USP4 and USP15 expression to lung cancer survival

Overall survival in patients with adenocarcinoma of the lung

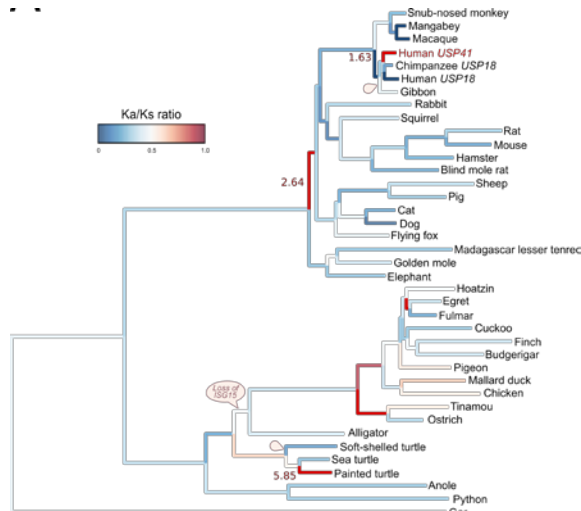


How to we apply evolutionary logic to infer protein function?



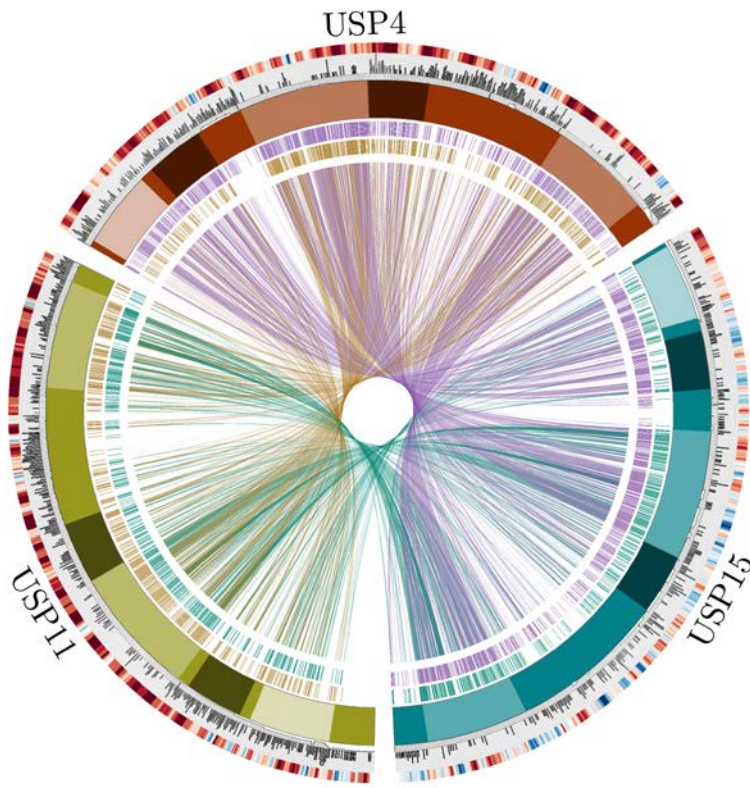
wikipedia.org

Sequence alignment and phylogeny-building

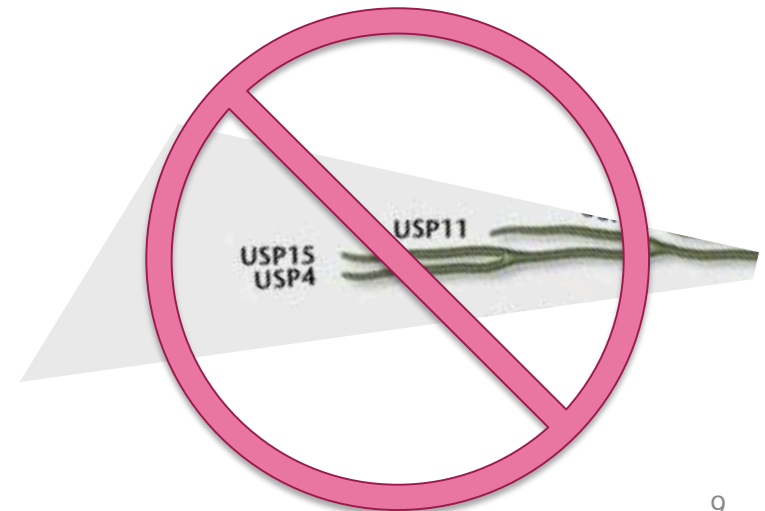
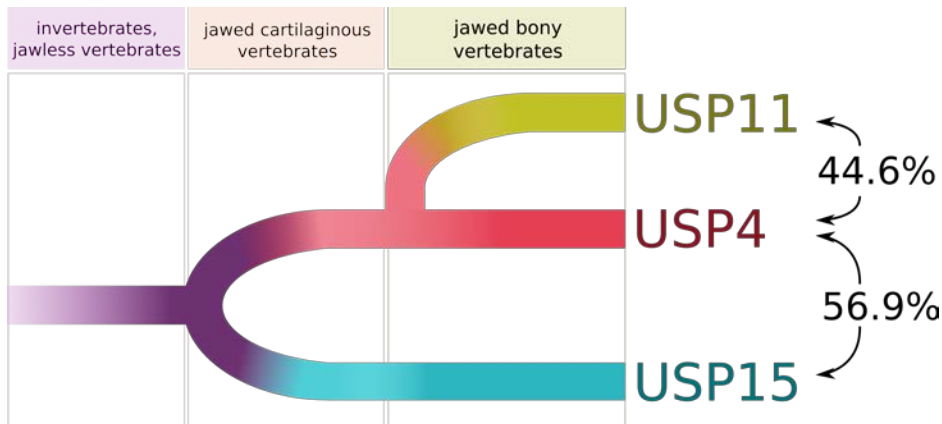


gi	ref	XP_006640391.1	ITVKIMTGKCSVDLEPSATVTELKQEVKRLAGIPETSQKLGINGGGHITVLEDHRT
gi 573905340	ref	XP_006640391.1	ITVKIMTGKCSVDLEPSATVTELKQEVKRLAGIPETSQKLGINGGGHITVLEDHRT
gi 318101506	ref	NP_001188203.1	LIITILNGQTHNVHVPNATVGEIKKEIAFRFKARPSQKLSITNGQIELDQDQRT
gi 323510630	ref	NP_001191098.1	ITVKILGGDVKRLEVSGDATVGIILKQVISOYFNVEFTFKOKLSAENQORISLEDESRT
gi 432874027	ref	XP_004072436.1	ITIVMLNGSTRITLVFPEDTVGSLKAQIETEMGVFAATQRLVVDNQRVTHSEDESRT
gi 118140098	gb	ABK63480.1	INIKMLNGTVHTLTVFPEDTVGSLKNIHGFQGEPPFKQKRLVVDNQRIDLSDDERT
gi 610391014	gb	AHW76805.1	IIIIIMLNGSTRITLVFPEDTVGSLKQVIOQKQGVFAERQRLVVDNQRIDLSDDERT
gi 262318083	dbj	BAI48419.1	IIIIIMLNGSTRITLVFPEDTVGSLKQVIOQKQGVFAERQRLVVDNQRIDLSDDERT
gi 658968045	ref	XP_008416907.1	ITVKMLDGTSHITLVFPEDTVGSLKMLQOKLQVFAARQRLVVDNQRIDLSDDERT
gi 908503562	ref	XP_013130844.1	ITITML-GESEHTLVFPEDTVGSLKIRIOEKLGVMHQRQRLVVDNQRIDLSDDERT
gi 305855021	dbj	BAJ16365.1	ITVKMLEGTECTLRVFPEDTVGSLKIRIOEKLGVMHQRQRLVVDNQRIDLSDDERT
gi 831277645	ref	XP_012669709.1	IQIKFLNGHTVPLAVSSITVGLDKKRIQKESDVTAPQOKLS-NNGINFSNDSST
gi 899144121	ref	NP_001297819.1	LTITILNGDSVPLTVFPHTVGSILSLIHOTLGAVTSTQRLSGVNGNNSLNDSEKT
anole_ISG15			LHFKMLTGEVHTLTVSPNHTM-EVKVLLERKMGCRRTYHOKIAAEAGSGID-LRDASS
sea_turtle_ISG15			LSVKLLTGMHSLVETSTMTVSAFKAQIAKKTGVSPVQOKLACONGA-YVELRDGSR
python_ISG15			LSVKLLTGEIHSVDSSACRTVDFKIQVGRKTVGPVQOKLACONSTHIN-LQDSSA
mouse_ISG15			LKVKMLGGNDFLVSTNMTVSELKKQIAQKIGVPAFQORLA-HQTA-VLQDGLT
rat_ISG15			LKVKMLGGKEFLVSTNMTVSELKKQIAQKIGVPAFQORLA-HQSG-EMLDQGVFA
bat_ISG15			LKVKML-SKEFQVPMRDSMLSELKQITOKTQVPAFQORLLVQSEN-AVLQDGVFP
rabbit_ISG15			LKVKMLSGKEIFLSESLTASELKRRIQKIGVPAFQORLLAVHPGG-AALIEEVP
tree_shrew_ISG15			LKVKMLGGQEFMVPISEMLAMELKQITKKIGVPAFQORLLATHPSG-TVLQDRVP
macaque_ISG15			LKVKMLGGNEFQVSLSEMSVSELKAKIAQKIGVPAFQORLLAVHPGG-ATLQDRVP
gorilla_ISG15			LKVKMLAGNEFQVSLSEMSVSELKAKIAQKIGVPAFQORLLAVHPGG-VALKDVFP
pan_paniscus_ISG15			LKVKMLAGNEFQVSLSEMSVSELKAKIAQKIGVPAFQORLLAVHPGG-VALQDRVP
human_ISG15			LKVKMLAGNEFQVSLSEMSVSELKAKIAQKIGVPAFQORLLAVHPGG-VALQDRVP
cow_ISG15			LTVKMLGQOEILVFLRDGMTVSELKQFIAQKINVPFAQORLLATHPSG-EVLQGVFP
platypus			QKVKMLNGKEFLVFLREMPVSELKQOIAQKIGVPAFQORLLAVHPAN-EVLQDGLP
pig_ISG15			LKVKMLGGKEILVFLRDGMTVSELKQOIAQKIGVPAFQORLLATHPSG-EVLQGVFP
dog_ISG15			LTVKMLGGEEFLVFLRDGMTVSELKQOIAQKIGVPAFQORLLATHPSG-TVLQDGLS
cat_ISG15			LKVKMLGGEEFLVFLRDGMTVSELKQOIAQKIGVPAFQORLLATHPSG-TVLQDGVFP
opossum_ISG15			LPLKVKMITGKEFTVNAQINMTVLNKKKIAAVKVPFAFQOKLA-TVGG-ELLINLNL
tasmanian_devil_ISG15			VMIILRVKMITGKEFSVKVEDNITVLDLKKRISQVGSFAHTQRLV-TEGG-EVLINLKL

- 1 Genealogy and radiation of **paralogs** USP4, 11, 15
- 2 Selection for USP4 **alternative isoforms**
- 3 Genealogy and radiation of **all deubiquitinases**



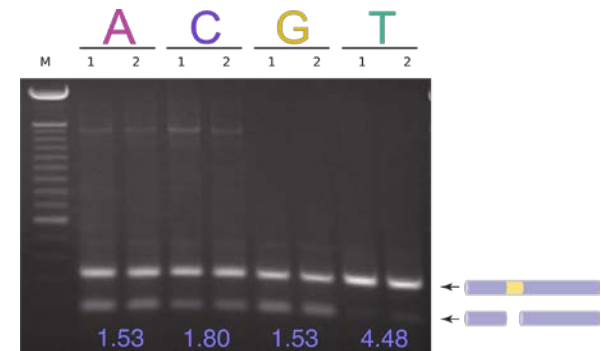
USP11 was derived from USP4 during a **small-scale duplication** event, which occurred after the whole genome duplication that produced USP4 & USP15.



In USP4, splice site nucleotides that favour alternative splicing are evolutionarily conserved...

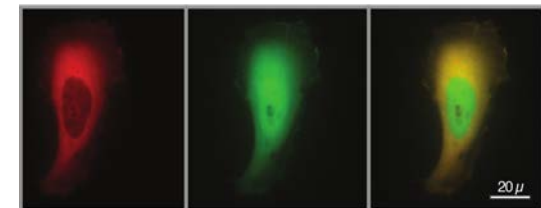


... mutating these sites changes splicing proclivity...



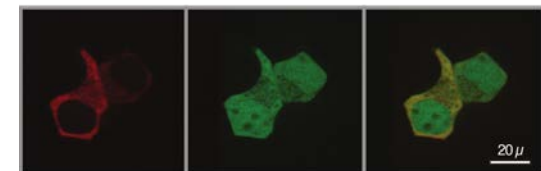
... the long and short isoforms have distinct subcellular localizations!

HeLa cell line

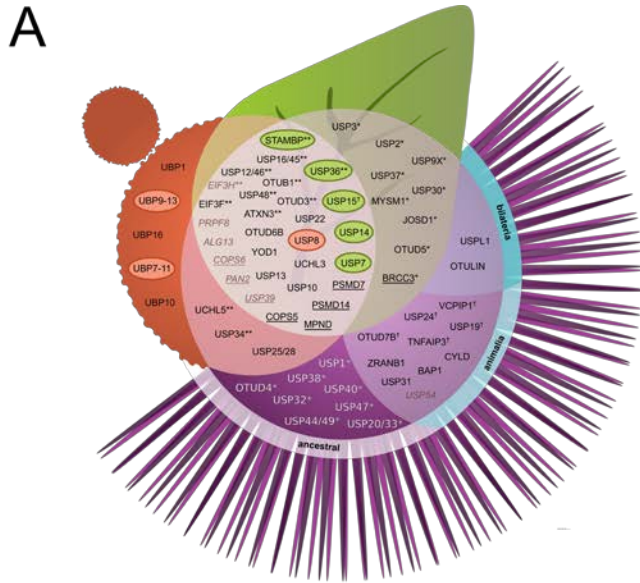


exon 7 retained exon 7 skipped merged image

293T cell line



exon 7 retained exon 7 skipped merged image



Expansion

from **29** DUBs in the eukaryotic ancestor
to **93** in the human genome

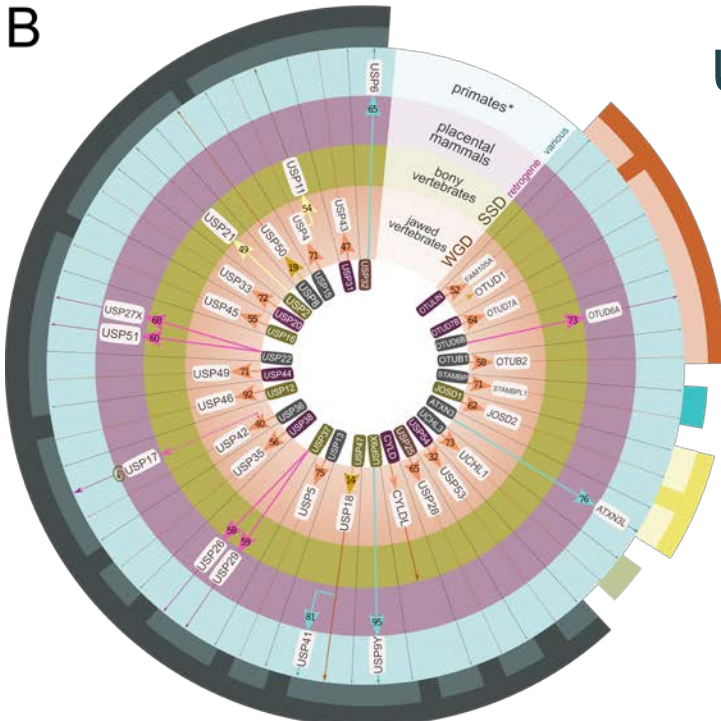
Radiation

e.g. **USP4** interacts with **CtIP** in DSB repair
while USP15 does not.

USP4 and CtIP were generated at the **same point in time** (whole genome duplication)
and they interact via **new domains**.

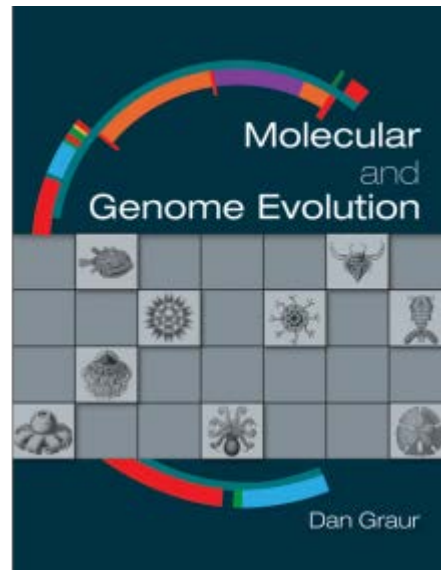
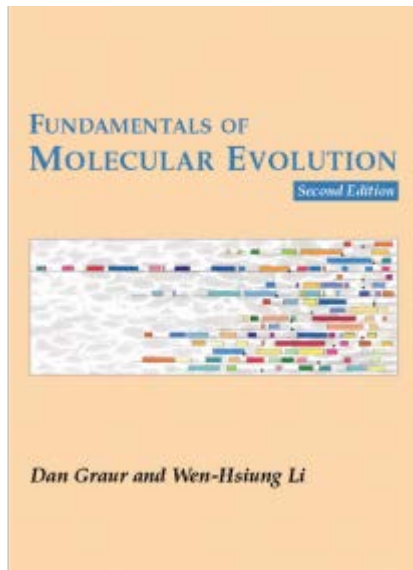
∴

USP4-CtIP represents
a new, specific interaction
(neofunctionalization)
that is maintained by natural selection



Take-away

- Molecular evolution analyses can be informative (and cost-effective) means to contextualize and guide wet lab research.
- Resources:
 - “Fundamentals of Molecular Evolution” by Graur & Li
 - “Molecular and Genome Evolution” by Graur



Thank you!

Gray lab
Dr. Doug Gray
Dr. Josée Coulombe

Xia lab
Dr. Xuhua Xia

Vanderhyden lab
David Cook