Linguistic Adequacy & the FoxP2 story

Evolution of Language '14
FGW

Jelle Zuidema
ILLC

Recap: scenarios

- Evolution of Humanness
  - Language-first
  - Intelligence-first
  - Massive modularity
- Evolution of Language
  - Gradualist, domain-specific scenario (Pinker)
  - Saltationist, domain-specific scenario (Chomsky)
  - Saltationist, domain-general scenario (Tomasello)
  - Gradualist, domain-general scenario (Deacon)

Recap

Can we assess the apriori evolutionary plausibility of different scenarios?

Recap: Components of evolutionary explanations

- Heritability & variation
- Strategy set
- Fitness
- Path of ever increasing fitness
  - Frequency-dependent fitness: solutions to the problems of coordination and cooperation (reciprocity, kin selection)

Recap: Limited time argument
(against gradualism)

- Implausible to assume very many *selected* genes for each uniquely human trait

Recap: Social traits argument
(against domain-specific gradualism)

- Most uniquely human traits are 'unusual', social traits in evolution because their benefits are *for* or *dependent on* the social group.

- E.g., evolution of language, communication, music, cooperativity, social cognition etc. all pose coordination and altruism problems:
  - Kin selection / Social evolution theory
  - Frequency dependent selection

- Require unusual circumstances; implausible to simply assume a prolonged selection regime favoring social traits

"No miracles" argument
(against saltationism)

- Scenarios should not give a major explanatory role to unknown properties of genes or brains or unknown laws of physics
  - e.g., a "macromutation" that out of a sudden creates a Universal Grammar (Bickerton'90)
  - e.g., "we know very little about what happens when 10^{10} neurons are crammed into something the size of basketball" - Chomsky'75
  - e.g., intricacies of morphosyntax might be explainable from cooperativity, social interactions & embodiment...

Requirements for plausible scenarios

- Explain how such a radical new phenotype can be based on relatively few genetic changes
  - (1) Common causes
  - (2) Hidden potential
- Explain how the unusual circumstances needed for the evolution of social traits can be sustained
  - (3) Self-enforcing dynamic

Language & cognition

- Reasoning: logic <-> language (not, and, or, if, then, all, every, some, X is Y, ...)
- Planning: hierarchical plans <-> hierarchical phrase-structure
- Theory of mind: intentional embedding <-> sentential embedding
- Mathematics: number words, context-free syntax of algebra
- Music: pitch, rhythm, phrasal structure, cultural transmission
- Consciousness: inner voice

Can we assess the linguistic adequacy of different scenarios?

Arguments

- Argument from Personal Incredulity (?)
- Argument from Authority (?)
- "No intermediate language" argument (against gradualist, domain-specific scenarios)
  - Across human individuals
    - But: SLI...
  - Across primate species
- Empirical adequacy arguments

Empirical Adequacy: Language Universals

- Consensus about universals at the level of 'design features', controversies about all specifics
  - Phonetics
  - Phonology
  - Lexicon
  - Morphosyntax
  - Semantics
  - Pragmatics

Empirical Adequacy: Evidence from language acquisition

- Chomsky (1959): Review of Skinner's Verbal Behavior;
- Chomsky 1981 – Principles & Parameters

vs.

- Connectionists (Bates, Elman, Seidenberg): general learning mechanism
- Usage-based/Construction grammar (Croft, Tomasello)

Contra nurture: children *resist* corrections

- Child: Nobody don't like me.
- Parent: No, say "nobody likes me."
- Child: Nobody don't like me.
- ...
- (Eight repetitions of this dialogue)
- …
- Parent: No, now listen carefully; say "nobody likes me."
- Child: Oh! Nobody don't likes me.
  - (David McNeill, 1970)

Slides from Jan de Jong

Contra nurture: Brown and Hanlon (1970)

- parents correct meaning, *not* form
  - Father: *Where is that big piece of paper I gave you yesterday?* Child: *Remember? I writed on it.* Father: *Oh, that's right*
- when correction was given, it was not picked up by the child
- 'no negative evidence' and 'poverty of the stimulus' strengthened the nativist position

Slides from Jan de Jong

What (other) evidence for nativism?

- Developmental → Grammar acquired effortlessly and systematically
- **Dissociations → Language and general intelligence dissociate in Williams Syndrome (see Karmiloff-Smith for evidence to the contrary) and in Specific Language Impairment**
- **Genetic → Grammatical impairment runs in families (Gopnik, Tomblin) Also cf. twin research (Bishop)**

The poverty of the stimulus

- "Any aspect of language that cannot be found in the child's linguistic environment, must be part of the innate equipment of the human brain" (Cook, 1983).
- The girl is lying
- Is the girl lying?
- The boy who plays piano is crying.
- Is the boy who plays piano crying?
- #Plays the boy who piano is crying?

(1)   a. a violin which this sonata is hard to play upon

          b. *a sonata which this violin is hard to play upon (Steedman'03)

(2)   a. Every acorn grew into an oak.

          b. Every oak grew out of an acorn.

          c. An oak grew out of every acorn.

          d. *An acorn grew into every oak. (Gruber, 1965)

(3)   a. a book which I hope I will write, and I fear that most people will burn without reading

          b. *Three mathematicians in ten derive a lemma and in a hundred prove completeness (Steedman'03)

- 1. Formalism necessary to represent natural language syntax is quite complex (trans-finite-state, and probably trans-contextfree);

-

-

-

- 1. Formalism necessary to represent natural language syntax is quite complex (trans-finite-state, and probably trans-contextfree);
- 2. There are many seemingly arbitrary features in the syntax of natural languages, each with many a priori plausible, but non-attested alternatives;
-
-

- 1. Formalism necessary to represent natural language syntax is quite complex (trans-finite-state, and probably trans-contextfree);
- 2. There are many seemingly arbitrary features in the syntax of natural languages, each with many a priori plausible, but non-attested alternatives;
- 3. Children nevertheless reliably arrive at the correct rules of grammar, based on relatively few example sentences and no negative feedback (the poverty of the stimulus): innate prior knowledge of language?;
-

- 1. Formalism necessary to represent natural language syntax is quite complex (trans-finite-state, and probably trans-contextfree);

- 2. There are many seemingly arbitrary features in the syntax of natural languages, each with many a priori plausible, but non-attested alternatives;

- 3. Children nevertheless reliably arrive at the correct rules of grammar, based on relatively few example sentences and no negative feedback (the poverty of the stimulus): innate prior knowledge of language?;

- 4. Languages differ greatly at first sight, but detailed analysis reveals many similarities between distinct languages in the underlying structure: a universal plan underlying all languages?

Debate continues

- Logic of the poverty of stimulus argument has been challenged (Pullum & Sholz'02; Zuidema'03)

- That the human learning apparatus is constrained in various ways is uncontroversial; the question is whether constraints are language-specific

- To argue for a domain-specific innate UG, it relies on personal incredulity: "I cannot imagine domain-general learning mechanisms to yield observed behavior"

The story of FoxP2

Genetic evidence to settle these debates?

- KE family (Hurst et al 1990, Gopnik 1990)
- Dispraxia (Vargha-Khadem 1995)
- FOXP2 (Kai et al 2001)
- Phylogeny (Enard et al 2002)
- Songbirds (Haesler et al 2004)

KE pedigree

○ female
□ male
● ■ affected
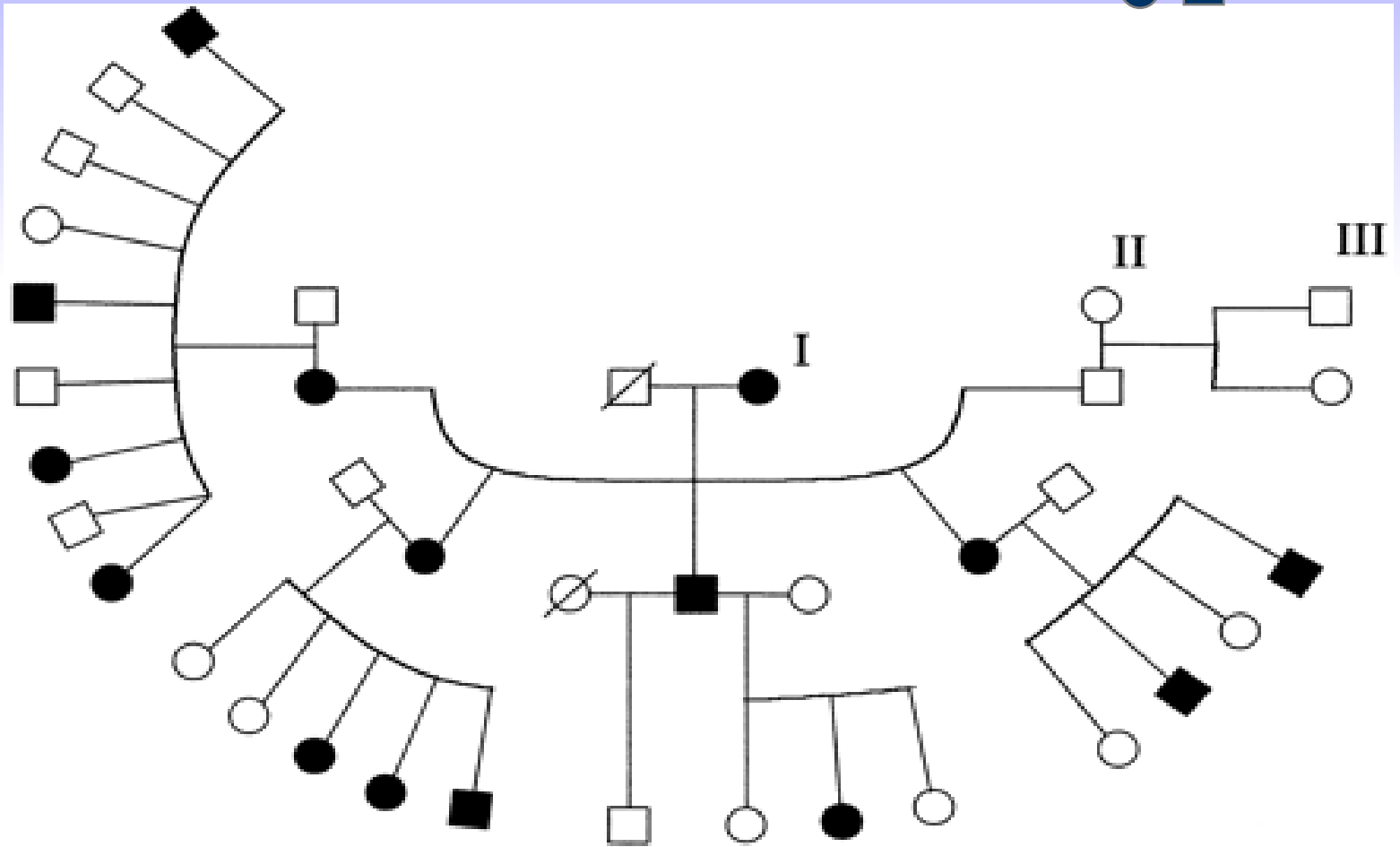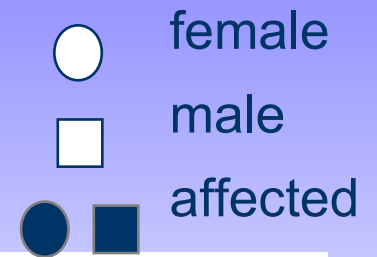
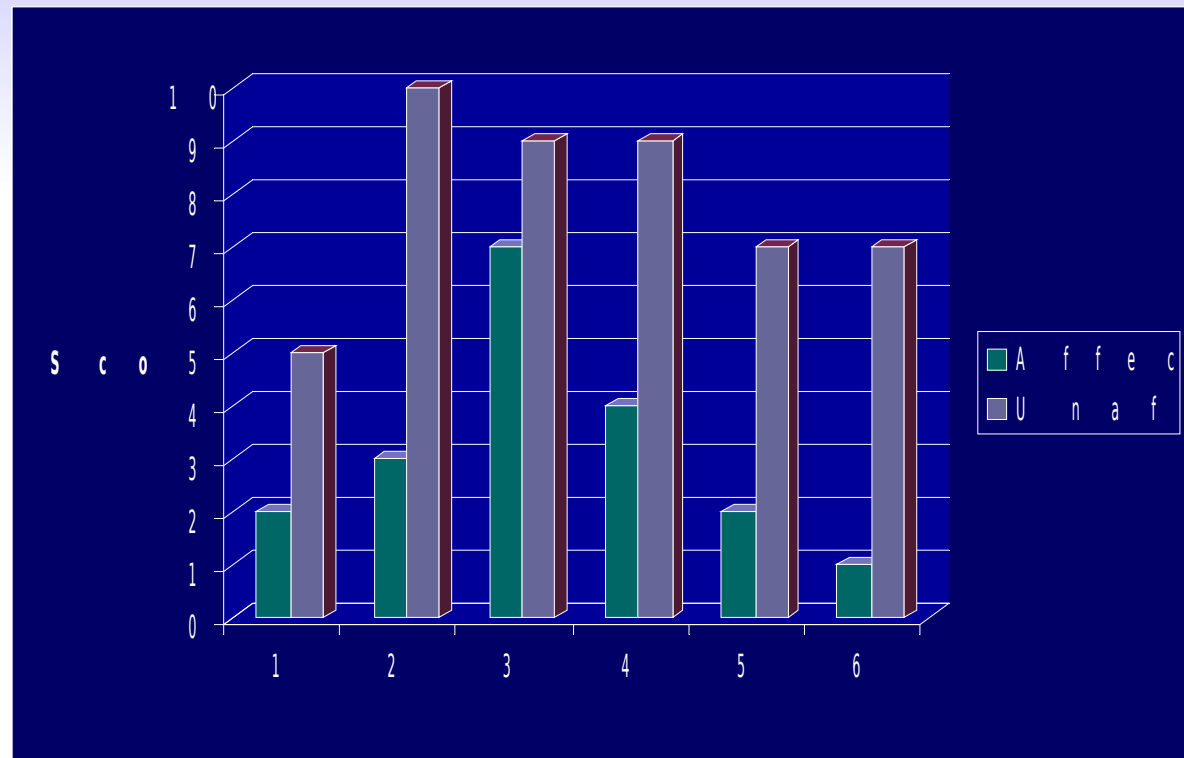Hurst et al (1990), Dev Med Child Neurol 32:347:355

Genetics: SLI Familial aggregation (FOXP2)

Gopnik & Crago
(*Cognition,* 1991):
specific impairment of
grammatical morphology
(e.g. number, tense)

Production of
expected tense
markings
{"Every day he walks 8
miles. Yesterday he
___ "

Slides from Jan de Jong

Specific language impairment (SLI) – a definition

- "children for whose non-normal language acquisition there is *no identifiable* physical or psychological basis. These individuals have normal hearing, intelligence within normal limits, an apparently intact neurological substrate, and no behavioural or emotional disorder. They *nevertheless* have persisting linguistic difficulties." (Fletcher, 1999 – italics added)
- *Exclusionary* definition 'ensures' *dissociation*

Slides from Jan de Jong

Part of a transcript – Kim (3;8) (Bowen, 1999)

- Mother: Here... (she hands him a smaller ball and Kim puts both on the truck]
- Kim: Two balls. Up on truck. *Truck going*. Go.
- Mother: That's better! That's better!
- Kim: Yeah. [Politely] *More truck*? Please
- Mother: Hey! How about this one Kimmy!
- [She offers him an "antique" bread truck he had not noticed before]
- Kim: Cool! [To me] *Where you hide truck*? More truck? More?

Slides from Jan de Jong

Colin, 6;8

- *MOT:those are the flats are they ?
- *CHI: yes people .
- *MOT:and are all those are all the people that live in the flats are they ?
- *CHI: yeah tall .
- *MOT:0 .
- *CHI: tall tall at the top .
- *MOT:that one lives on the top ?
- *MOT:oops !
- *CHI: yeah .
- *MOT:they're all falling over .
- *CHI: tall .
- *MOT:you have to put them on very carefully .
- *CHI: very carefully .
- *MOT:no the man won't stand on top of his head .
- *CHI: other mans man live there .
- *MOT:that man lives there does he ?
- *CHI: man lives here fat controller house .
- *MOT:he's the fat controller's house ?
- *CHI: yeah .

Slides from Jan de Jong

Vargha-Khadem *et al.* (1995)

- KE-family disorder is not speech or grammar-specific
- Defects in intellectual, linguistic and orofacial praxic function
- In particular, affected individuals are much worse at imitation of oral and facial movements
- Linguistic difficulties do constitute prominent part of of phenotype
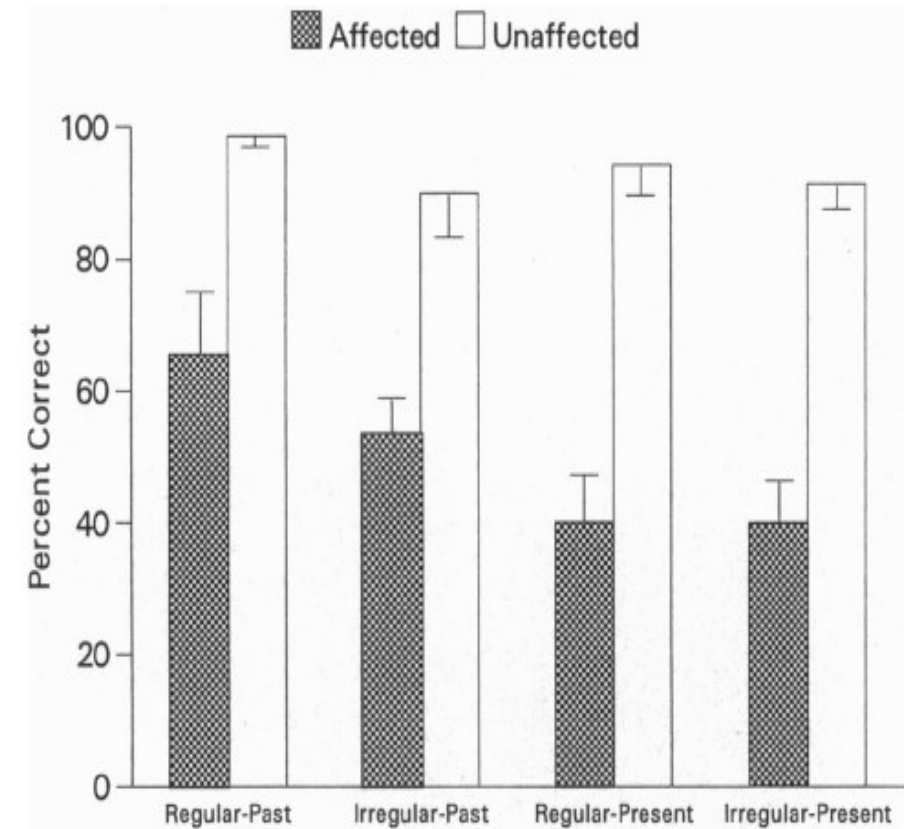 PNAS 92:930-933

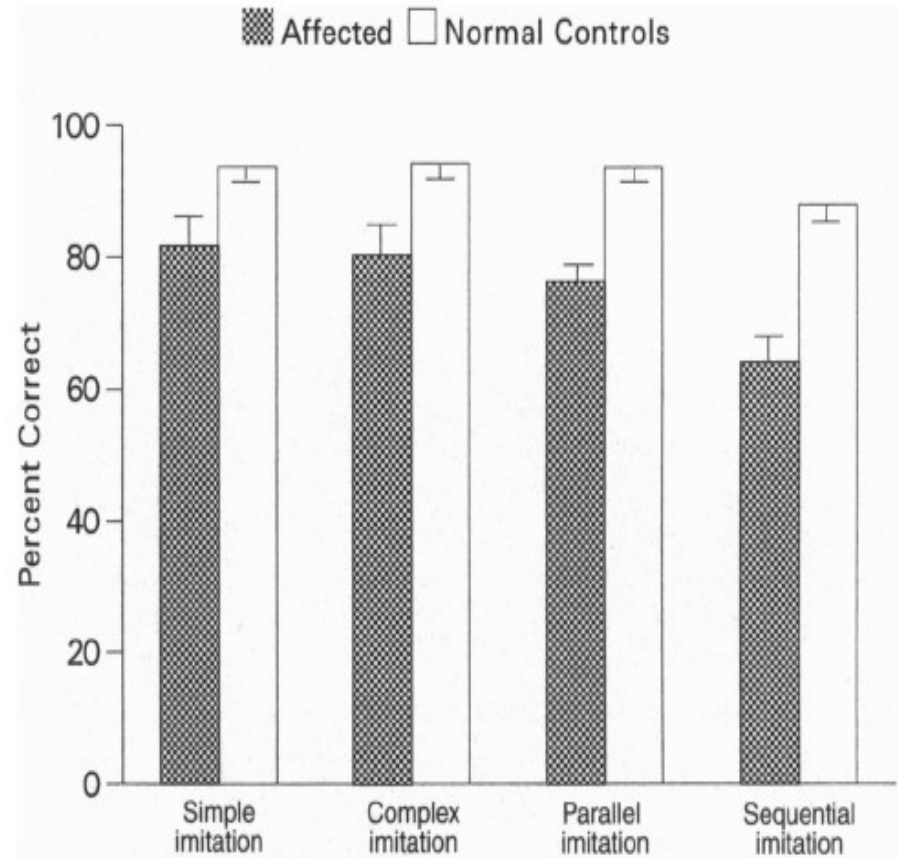FIG. 2. Production of tenses. Scores are means ± standard errors. See Table 2 for examples of test items.

FIG. 3. Imitation of oral and facial movements. Scores are means ± standard errors.
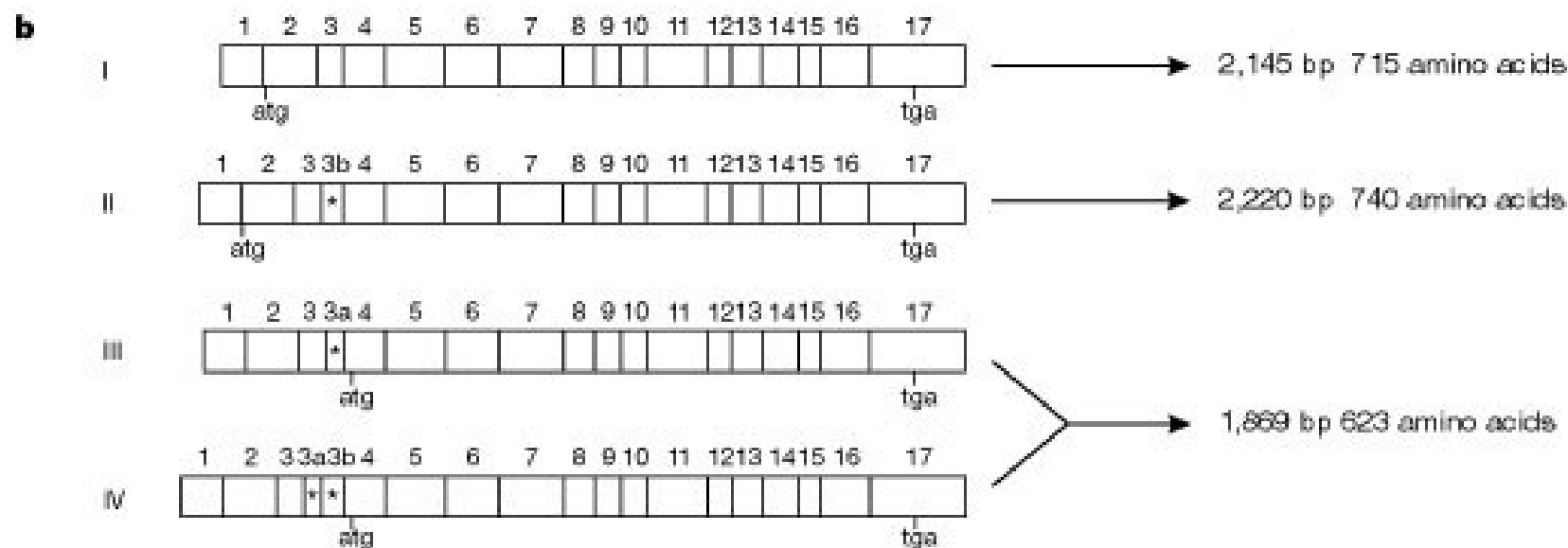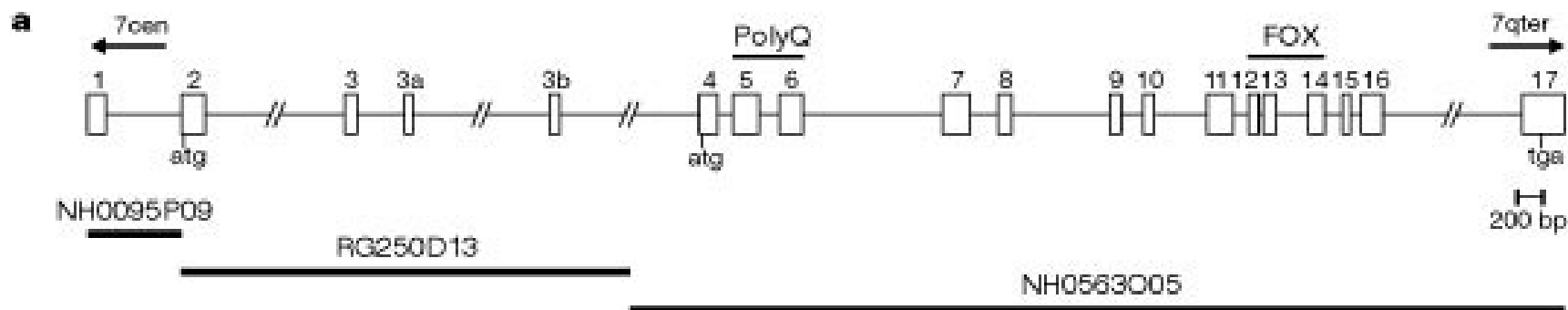
Vargha-Khadem *et al.* (1998)

- *Neural basis of an inherited speech and language disorder* (PNAS 95:12695-12700)
- Brain abnormalities in affected family members: structural/size (MRI), function/activity (PET)
- Bilateral reduction in size caudate nucleus
- Abnormal high activity in left caudate nucleus during speech tasks
- Broca's area smaller and overactivated

Fisher *et al.* (1998)

- *Nature Genetics* 18:168-170
- Linkage study
- Narrowed mutation to stretch on chromosome 7 (7q31)
- Named gene SPCH1

Lai *et al* (2001)

- *A forkhead-domain gene is mutated in a severe speech and language disorder* (Nature 413:519-523)
- Identified unrelated individual CS, with impairment and chromosomal translocation involving SPCH1 interval
- Identified point mutation in FOXP2: transcription factor containing forkhead DNA-binding domain (& polyglutamine tract)

a

7cen ←

1 2 3 3a 3b 4 5 6 PolyQ 7 8 9 10 11 12 13 FOX 14 15 16 17

atg   atg   tga

→ 7qter

NH0095P09

RG250D13

NH0563O05

200 bp

b

I

1 2 3 4 5 6 7 8 9 10 11 1213 1415 16 17

atg   tga

→ 2,145 bp 715 amino acids

II

1 2 3 3b 4 5 6 7 8 9 10 11 1213 1415 16 17

atg *   tga

→ 2,220 bp 740 amino acids

III

1 2 3 3a 4 5 6 7 8 9 10 11 1213 1415 16 17

atg *   tga

IV

1 2 3 3a 3b 4 5 6 7 8 9 10 11 1213 1415 16 17

atg * *   tga

→ 1,869 bp 623 amino acids

c

```
FOXP2  MMQESATETISNSSMNQNGMSTLSSQLDAGS-RDGRSSGDTSS-EVSTVELLHLQQQQ--ALQAARQLLLQQQ----TSGLKSPKSSDKQRPLQVPVSVA  92
       MMQES TET SN S  QNG    + L+ G  R+GRS+G+T + ++    +L H QQQQ  ALQ ARQLLLQQQ      SGLKSPK +DKQ  LQVPVSVA
FOXP1  MMQESGTETKSNGSAIQNGSGGSNHLLECGGLREGRSNGETPAVDIGAADLAHAQQQQQQQALQVARQLLLQQQQQQQVSGLKSPKRNDKQPALQVPVSVA  100

FOXP2  MMTPQVITPQQMQQILQQQVLSPQQLQALLQQQQAVMLQQQQLQEFYKKQQEQLHLQLLQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQH  192
       MMTPQVITPQQMQQILQQQVLSPQQLQ LLQQQQA+MLQQQQLQEFYKKQQEQL LQLLQQQ                                    H
FOXP1  MMTPQVITPQQMQQILQQQVLSPQQLQVLLQQQQALMLQQQQLQEFYKKQQEQLQLQLLQQQ-------------------------------------H  163

FOXP2  PGKQAKEQQQQQQQQQLAAQQLVFQQQLLQMQQLQQQHLLSLQRQGLISIPPGQAALPVQSLPQAGLSPAEIQQLWKEVTGVHSMEDN-GIKHGGLDL  291
        GKQ KE QQQ    +A QQL FQQQLLQMQQLQQQ HLLSLQRQGL++I PGQ ALP+Q L Q G+ P E+QQLWKEVT  H+ E+  G  H  LDL
FOXP1  AGKQPKEQQQ--------VATQQLAFQQQLLQMQQLQQQ-HLLSLQRQGLLTIQPGQPALPLQPLAQ-GMIPTELQQLWKEVTSAHTAEETTGNNHSSLDL  254
```

Enard et al (2002)
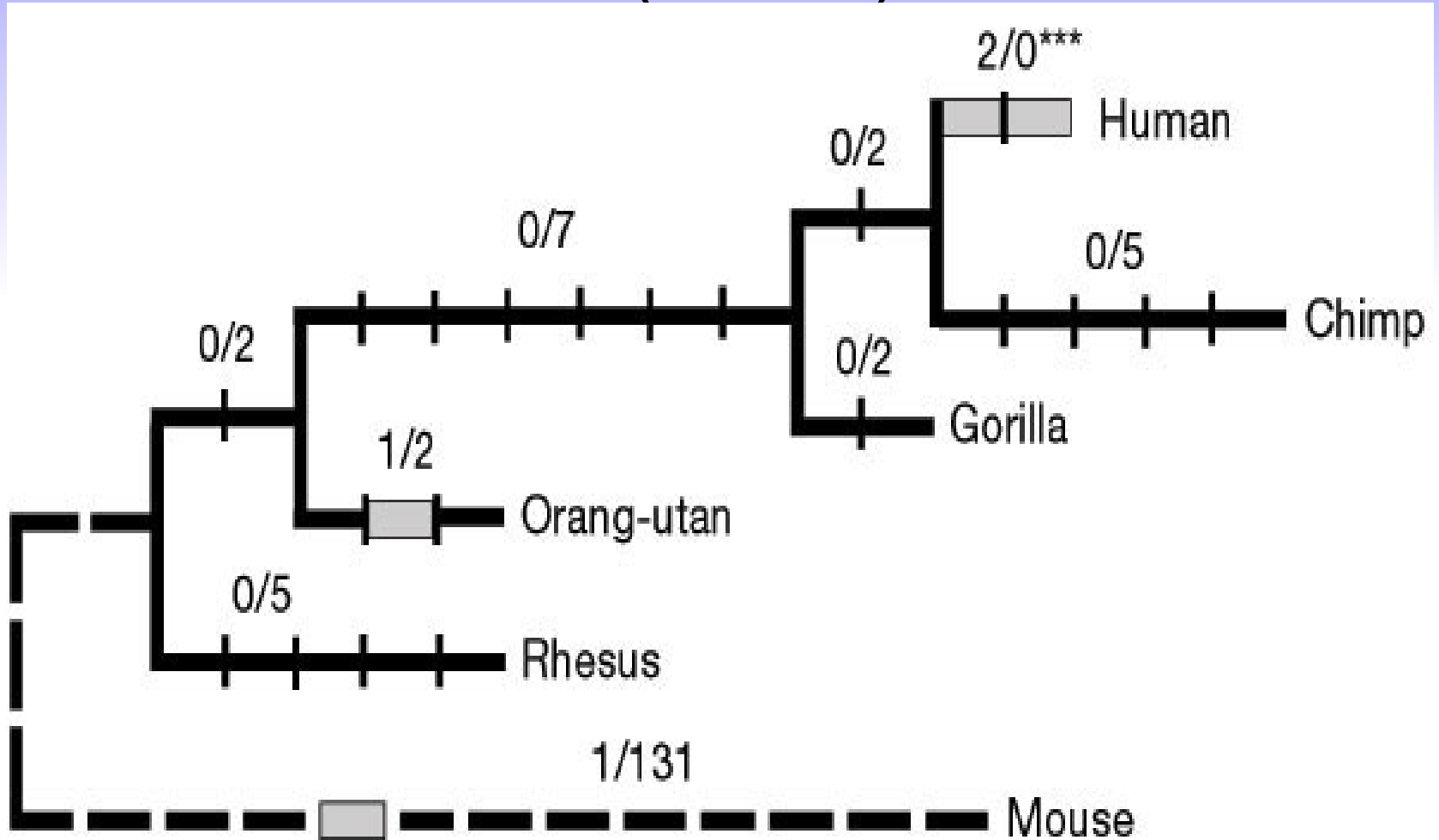
Enard et al (2002)

- *Molecular evolution of FOXP2, a gene involved in speech and language* (Nature:418:869-872)
- Evidence for selective sweep: reduced polymorphism in introns within FOXP2 and recombining loci
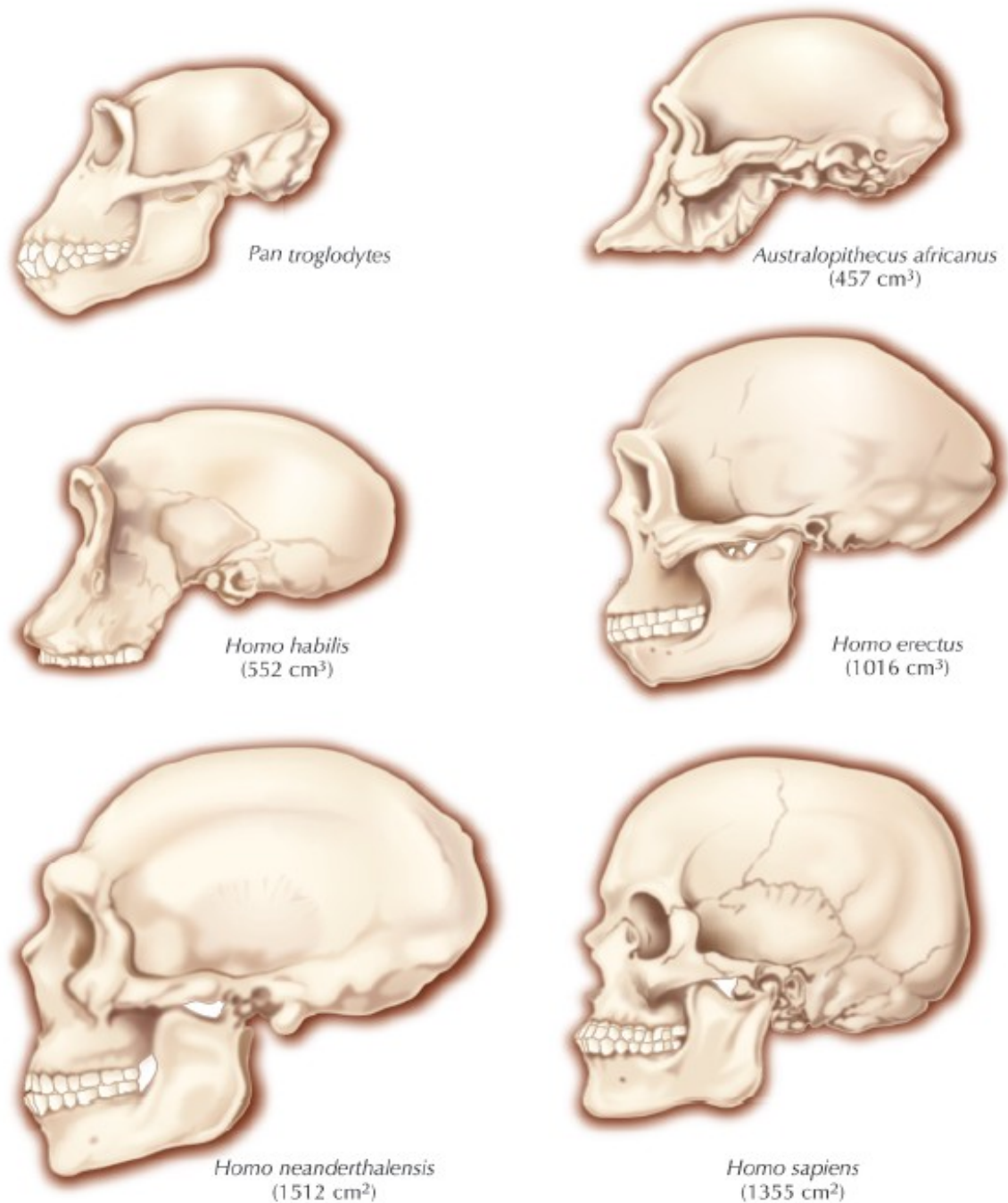- Estimated human-specific mutations happened between 10,000 and 100,000 years ago
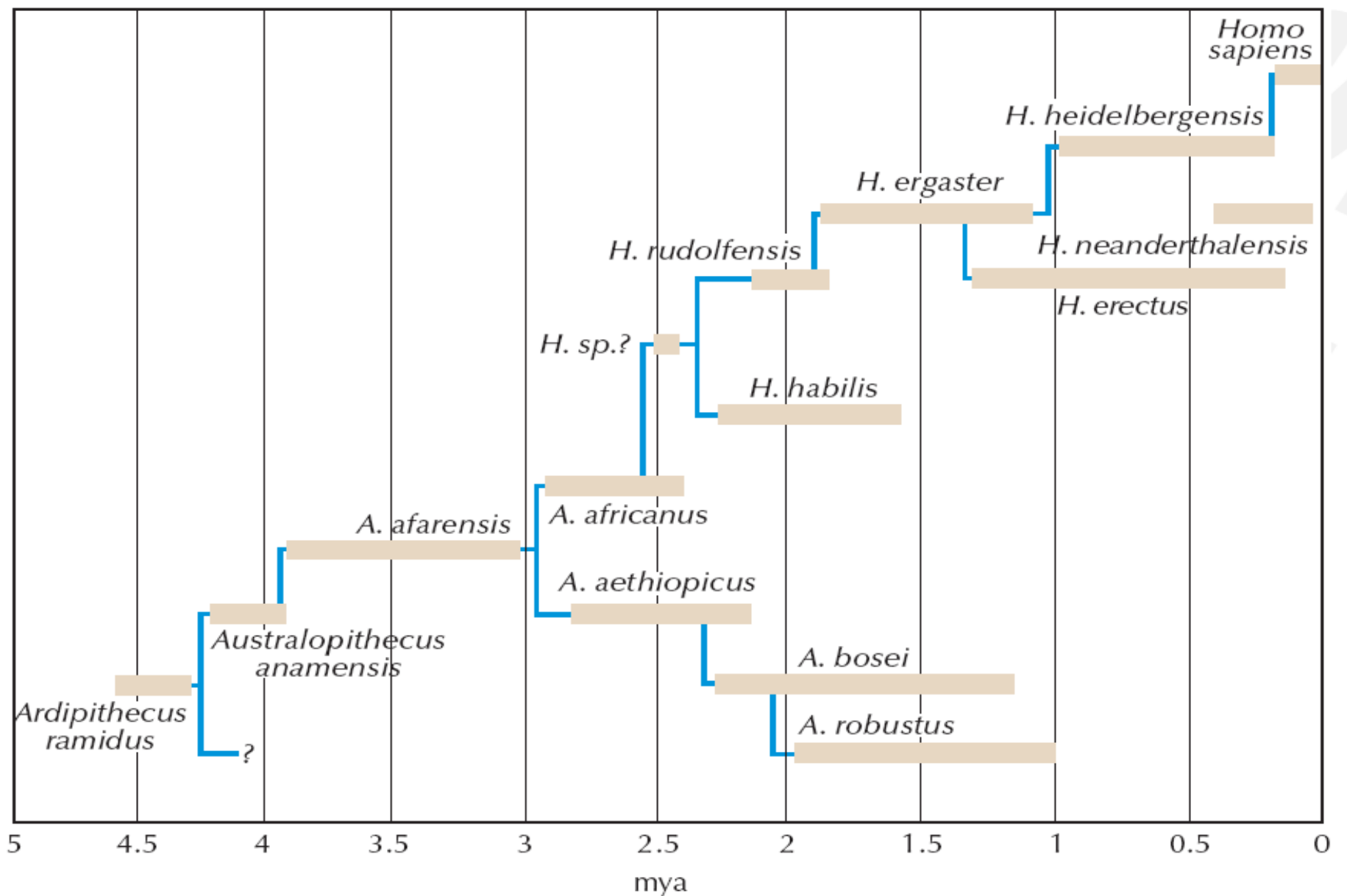
Pan troglodytes

Australopithecus africanus
(457 cm³)

Homo habilis
(552 cm³)

Homo erectus
(1016 cm³)

Homo neanderthalensis
(1512 cm²)

Homo sapiens
(1355 cm²)

FIGURE 25.9. Series of hominid skulls showing increasing size and brain volume.

URE 25.5. Timeline of major events of hominid evolution.

Zebra finches and other birds

- Haesler *et al.* (2004), FOXP2 expression in Avian Vocal Learners and Non-Learners (J. of Neuroscience 24(13):3164-3175)
- Area X expressed more FOXP2 than surrounding tissue at 35-50 days

Singing Mice

- Shu *et al.* (2005), Altered ultrasonic vocalization in mice with a disruption in the FOXP2 gene (PNAS)
- Disruption of both copies leads to severe motor impairment, premature death and absense of ultrasonic vocalizations in pups
- Disruption of single copy leads to modest developmental delay, but significant alteration in vocalizations; Purkinje cells particularly affected.
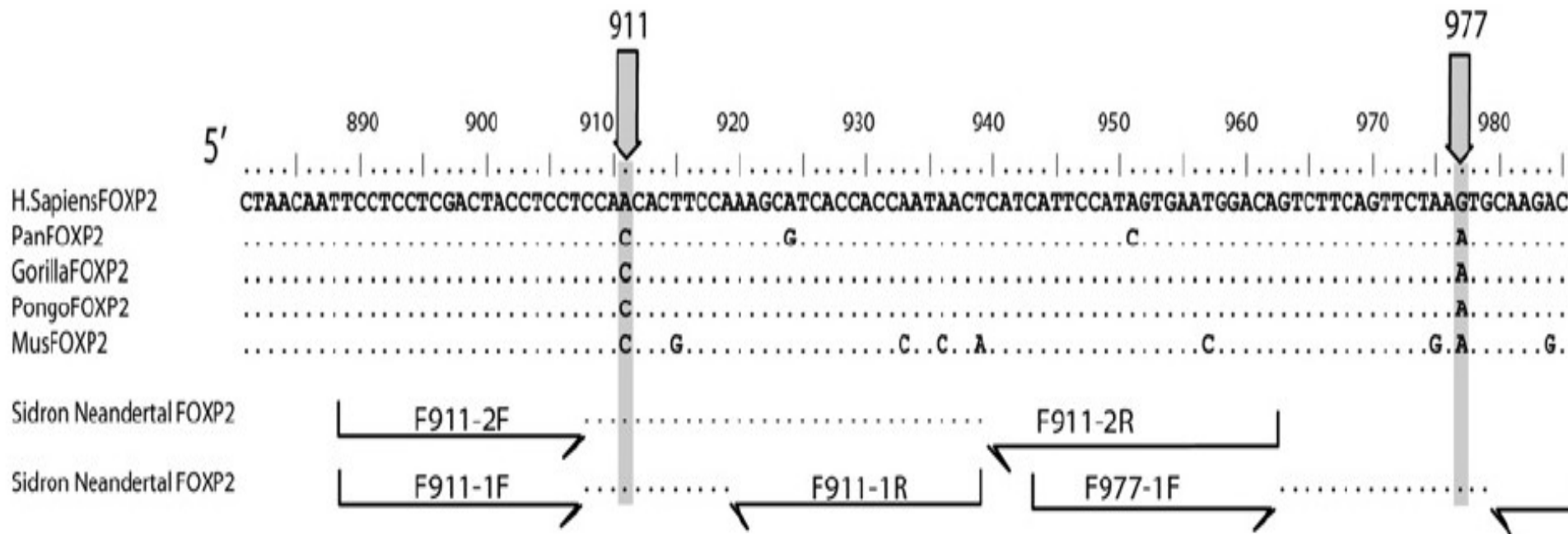
Neandertal genome

Figure 1. Sequence Alignment of Nucleotide Positions 880–1020 from the *FOXP2* Gene

The two nonsynonymous nucleotide substitutions on the human lineage are indicated by arrows. Identical posit dots. The three primer pairs used to retrieve the two substitutions from the El Sidrón Neandertals are indicate