When Social Networks Write Science History
Priority Disputes on Wikipedia

Presenter: Arno Simons

Co-authors: Wolfgang Kircheis, Marion Schmidt, Martin Potthast, Benno Stein

Workshop: "Framing Innovation in a Networked World" (2–3 September 2021)

How are CRISPR–related priority disputes enacted on Wikipedia?

Three bits of sociology

"Ready-made science" vs. "science-in-the-making" (Latour 1987)

Wikipedia = a setting for observing the construction of facts

"Retrospective accounts" (Deuten & Rip 2000)

Interpretations of the past to orient present action towards the future

The neglected sibling of "prospective accounts"

Studying "controversies" online (Venturini 2012)

Cartography: Recursive adjustment of exploration and representation

Distant reading + close reading

Three groups of actors...

1) Our own team

Arno Simons | DZHW | Political Science & STS

Wolfgang Kircheis | Leipzig U | Computer Science

Martin Potthast | U Leipzig | Digital Humanities

Marion Schmidt | DZHW | Science Studies & Bibliometrics

Benno Stein | BHU Weimar | Information Science

Wikipedia Science Analytics
https://osf.io/ygur7

2) Our "CRISPR heroes"

CRISPR pioneers (from left to right): George Church, Jennifer Doudna, Feng Zhang, and Emmanuelle Charpentier

© KEN RICHARDSON; © PETER RIGAUD C/O SHOTVIEW ARTISTS

SHARE

How the battle lines over CRISPR were drawn

By Jon Cohen | Feb. 15, 2017 , 3:30 PM

The U.S. Patent Trial and Appeal Board ruled today in favor of the Broad Institute in the high stakes battle over who will control the valuable intellectual property linked to CRISPR, the powerful genome-

Science's extensive COVID-19 coverage is free to all readers. To support our nonprofit science journalism, please **make a tax-deductible gift today.**
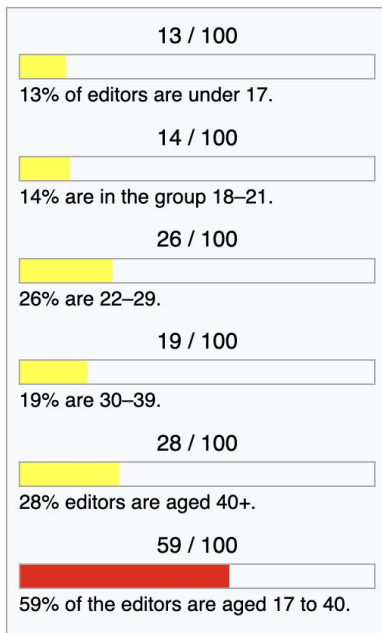
2) Our "CRISPR heroes"

3) Our Wikipedians

"Ppgardne"

"84.120.0.24"

"CatPath"

"PovilasTheGreat"

"Jytdog"

3) Our Wikipedians

Age distribution

13 / 100
13% of editors are under 17.

14 / 100
14% are in the group 18–21.

26 / 100
26% are 22–29.

19 / 100
19% are 30–39.

28 / 100
28% editors are aged 40+.

59 / 100
59% of the editors are aged 17 to 40.

Gender

84 / 100
The 2013 study **The Wikipedia Gender Gap Revisited** measured gender bias in survey completion and estimated that as of 2008, 84% of English Wikipedia editors were male. In the worldwide **Wikipedia Editor Survey 2011** of all the Wikipedias, 91% of respondents were male.

Nationality

The greatest number, or plurality, of editors (20%) reside in the United States, followed by Germany (12%) and Russia (7%). The only country not in Europe or North America in the top 10 is India (3%).

Language

76 / 100

49 / 100
Most users primarily edit (76%) and read (49%) the English Wikipedia, followed by the German Wikipedia at 20% and 12%, and the Spanish Wikipedia at 12% and 6% respectively. More than half (51%) of editors contribute in two or more languages.

Why contribute?

71 / 100
71% of the editors contribute because they like the idea of volunteering to share knowledge.

69 / 100
69% believe that information should be freely available.

63 / 100
63% pointed out that contributing is fun.

7 / 100
Only 7% edit Wikipedia for professional reasons.
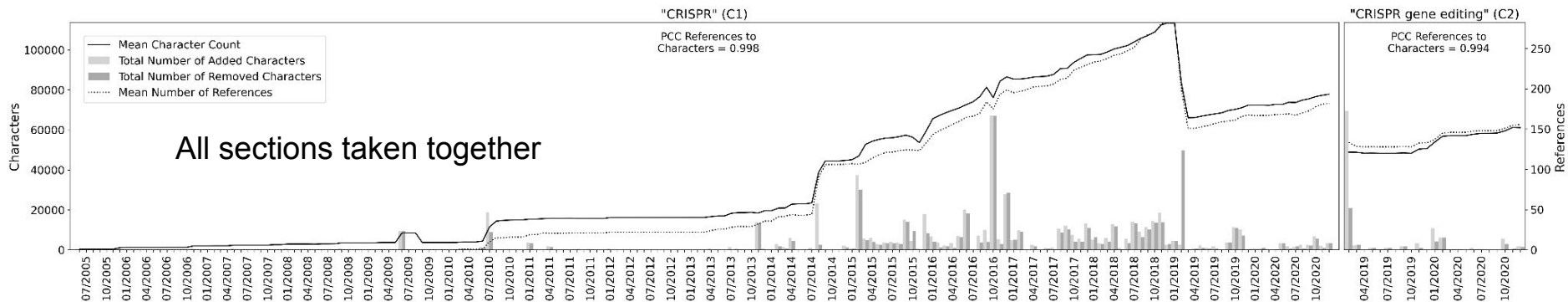
Distant and close readings

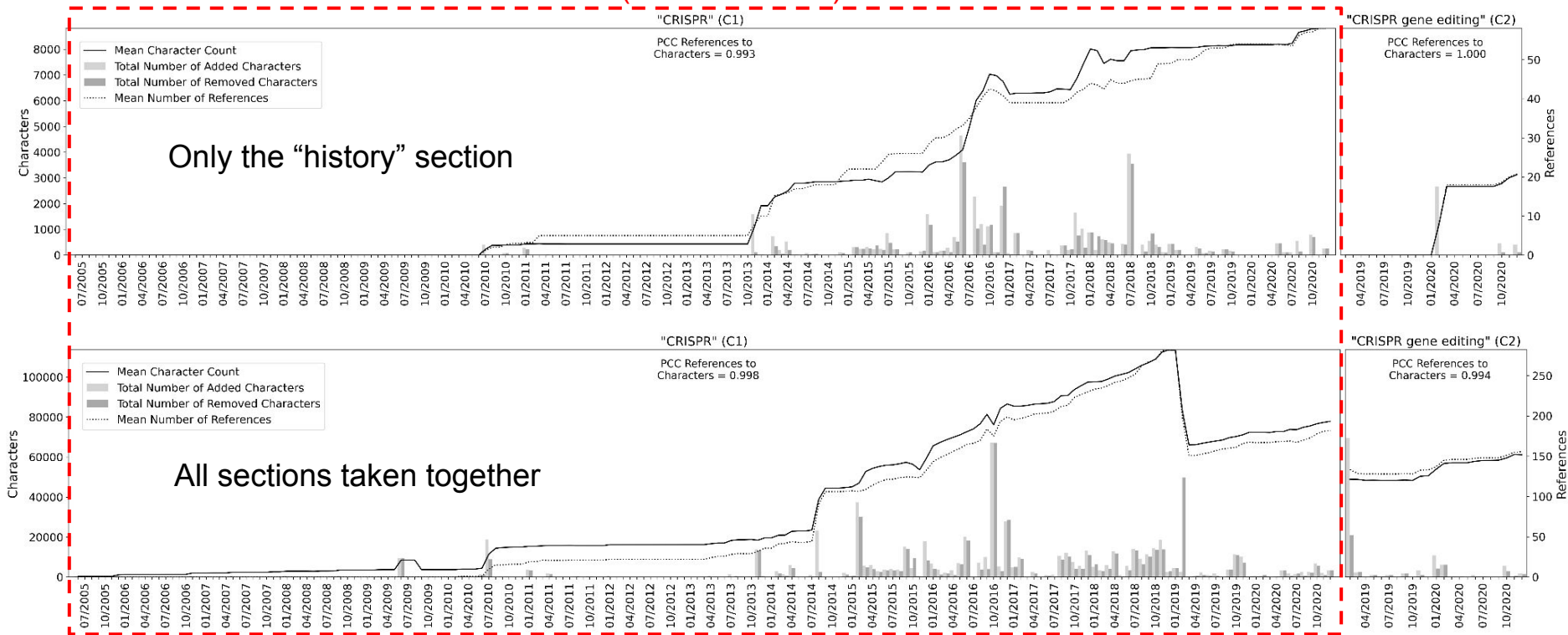# Inspecting the editing landscape from far

Only the "history" section
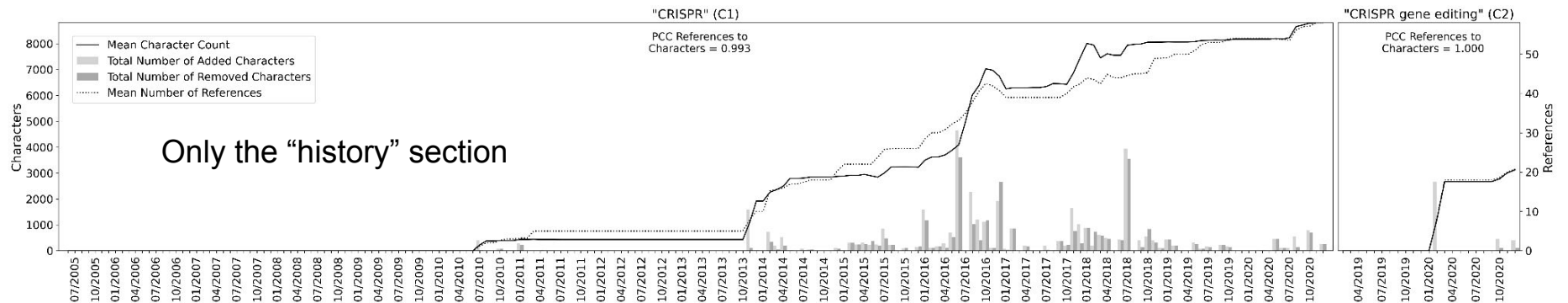
All sections taken together
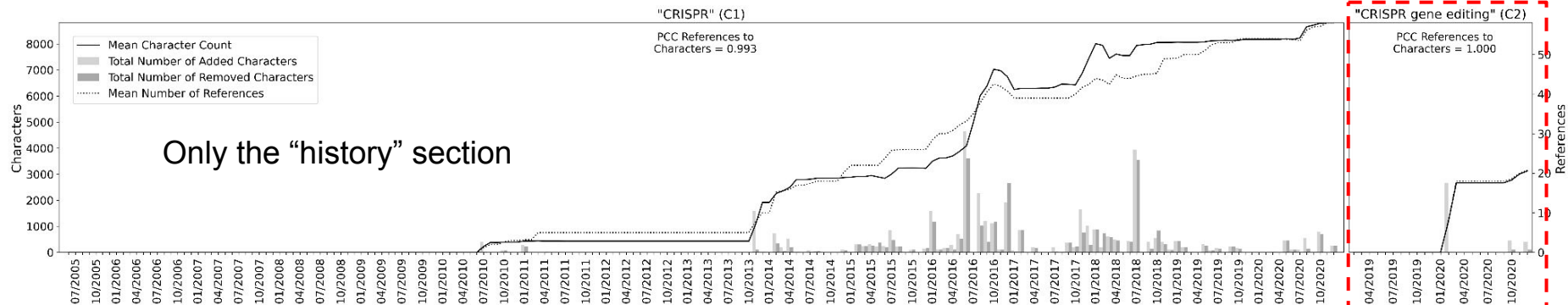
Inspecting the editing landscape from far
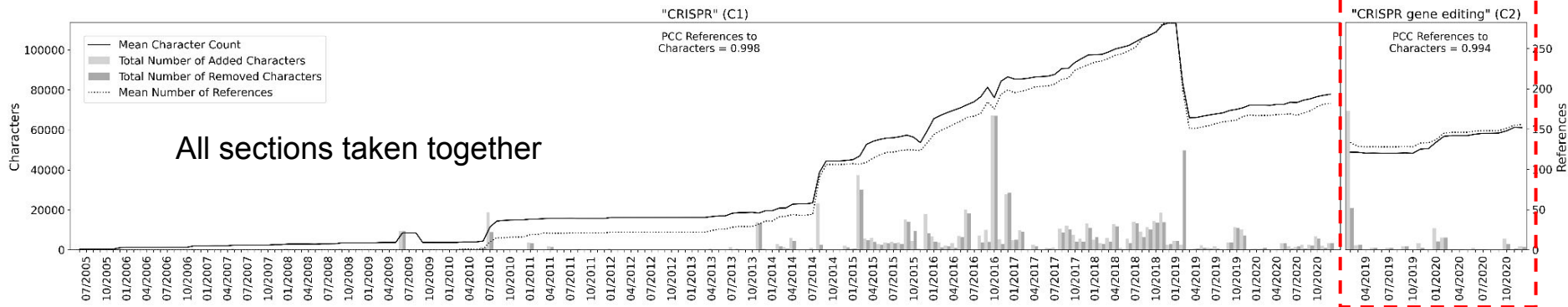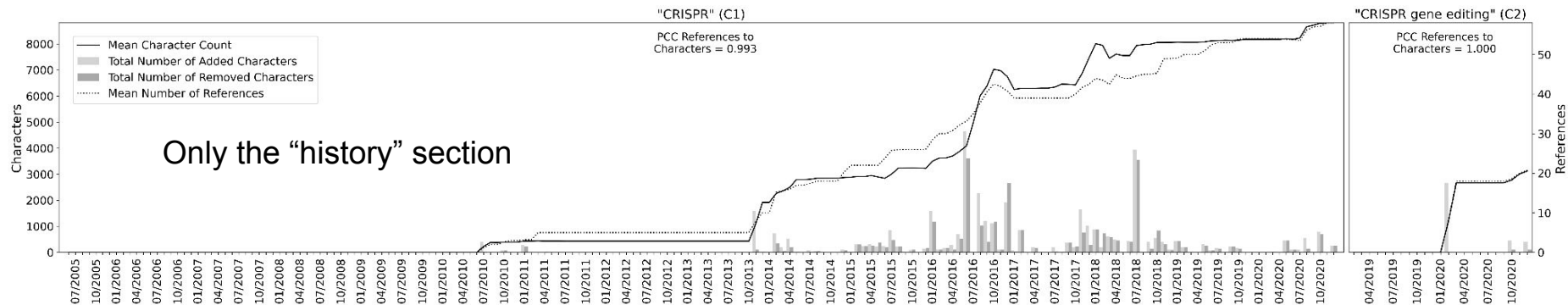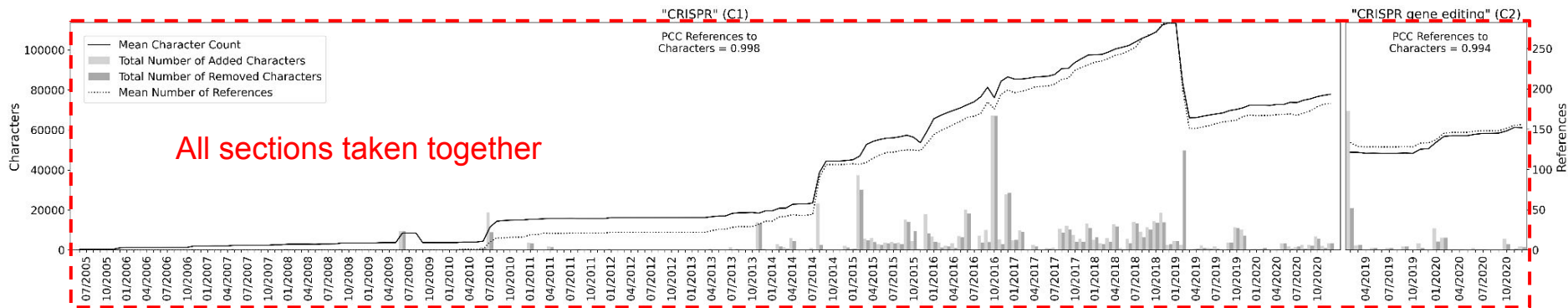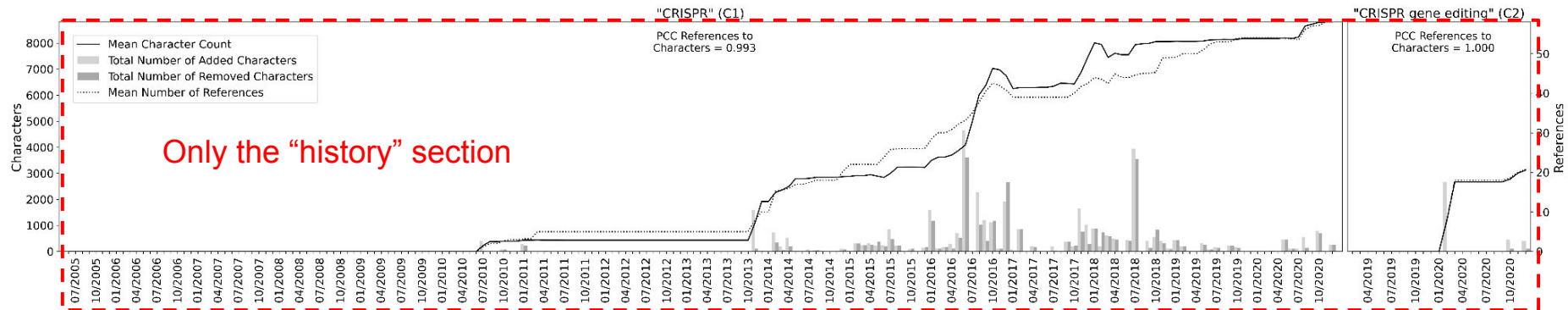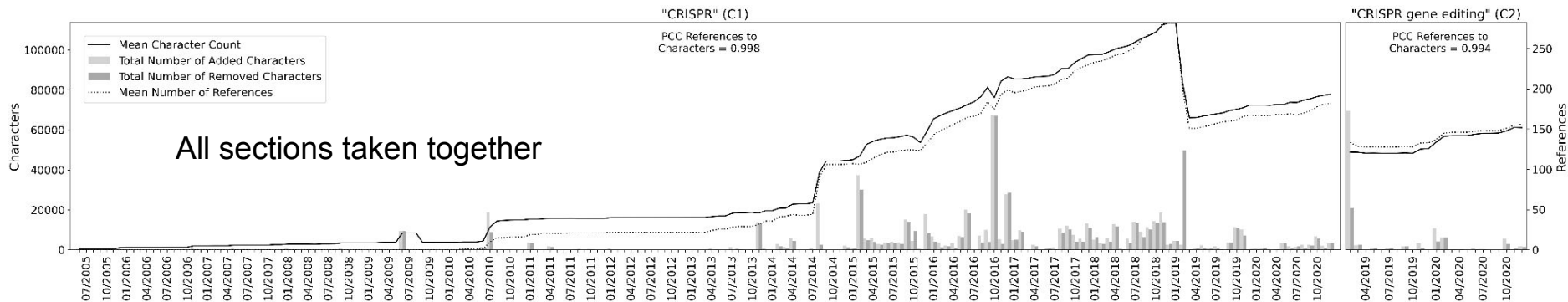

"CRISPR"
(2068 revisions)

Inspecting the editing landscape from far

Inspecting the editing landscape from far

Only the "history" section

All sections taken together
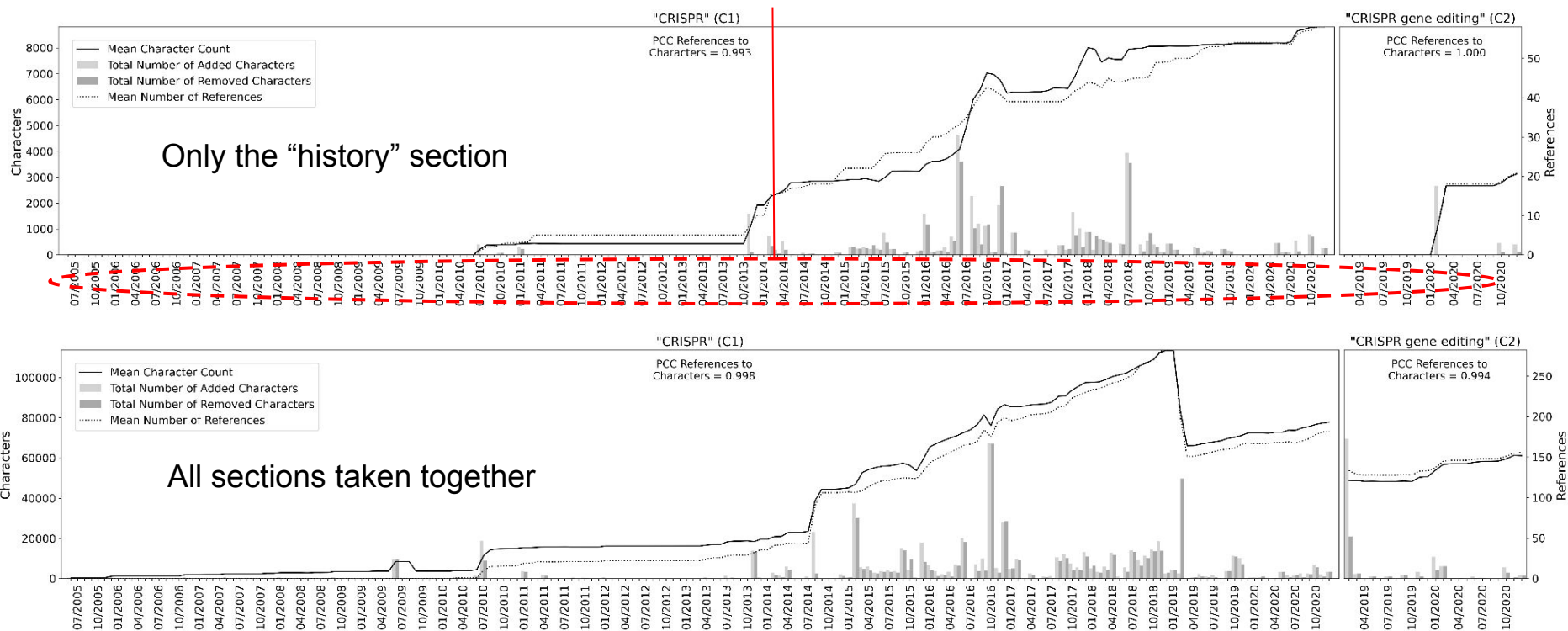
Inspecting the editing landscape from far

Inspecting the editing landscape from far
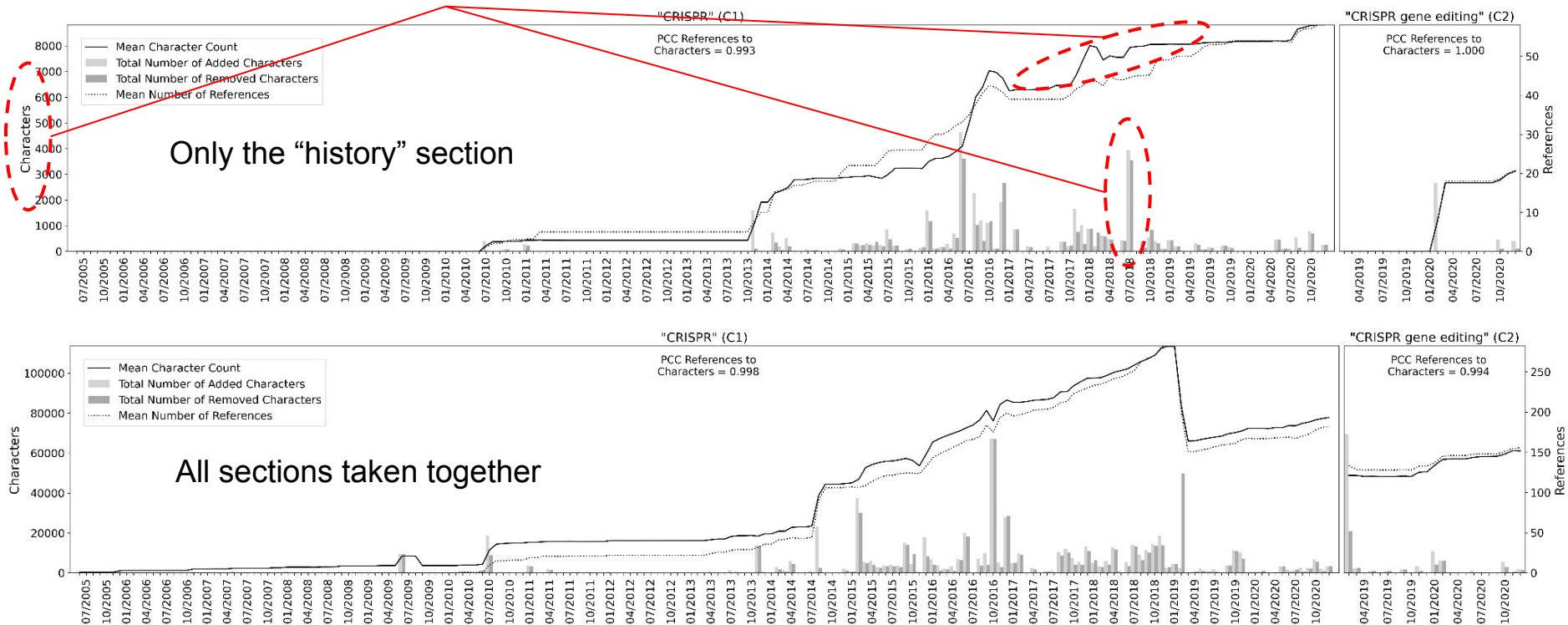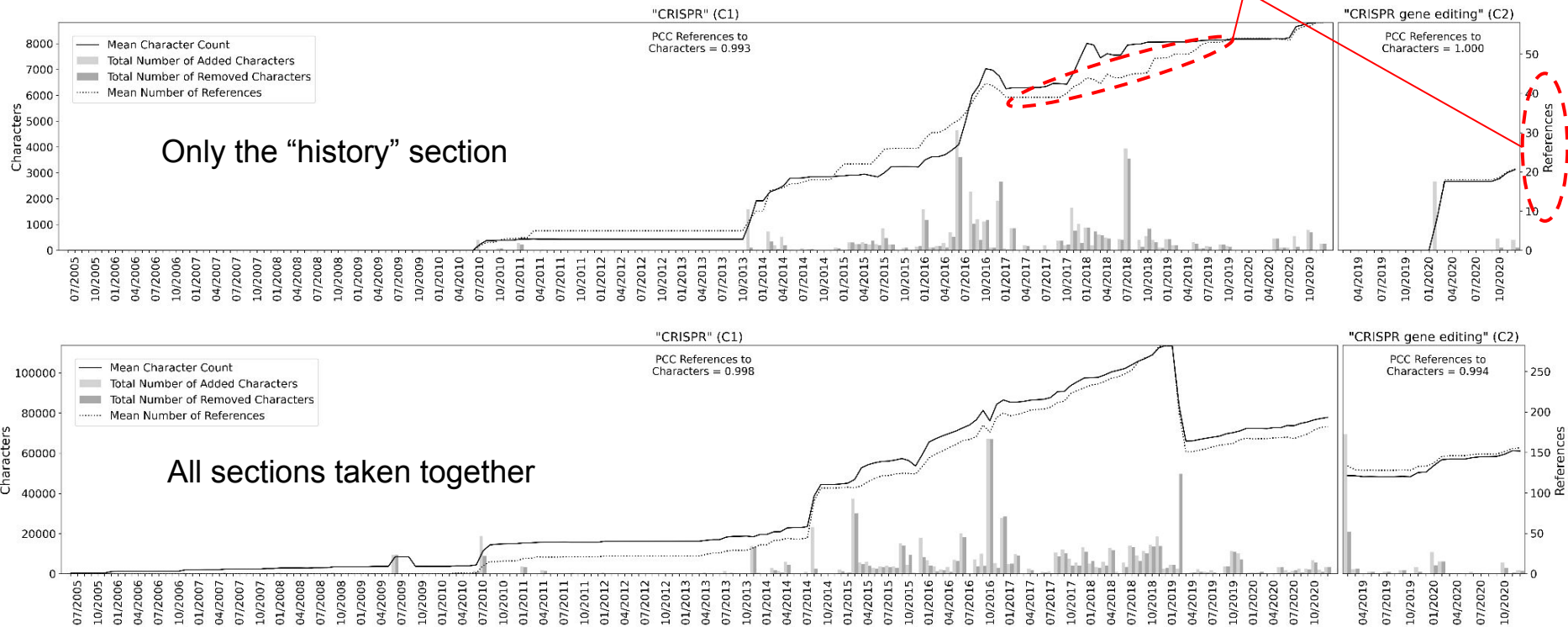

Only the "history" section

Inspecting the editing landscape from far

X-axis: Time (3-months slices)

Inspecting the editing landscape from far

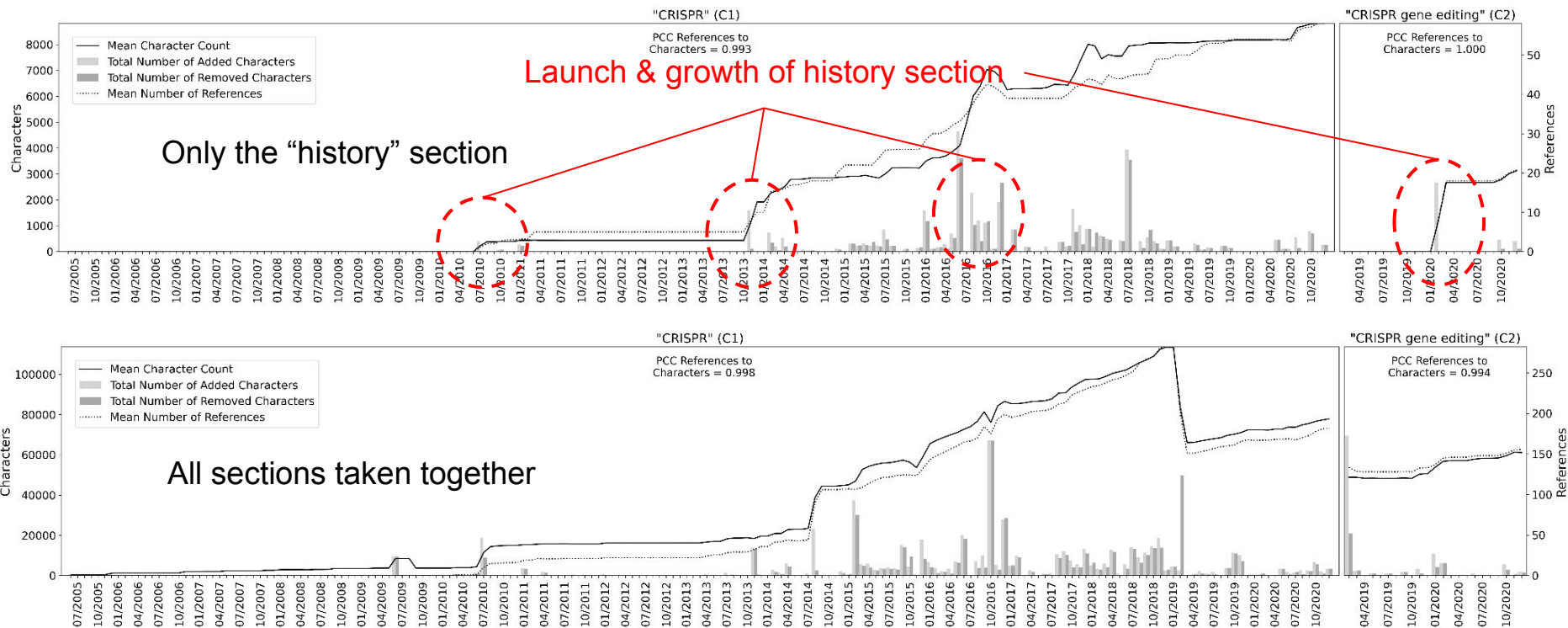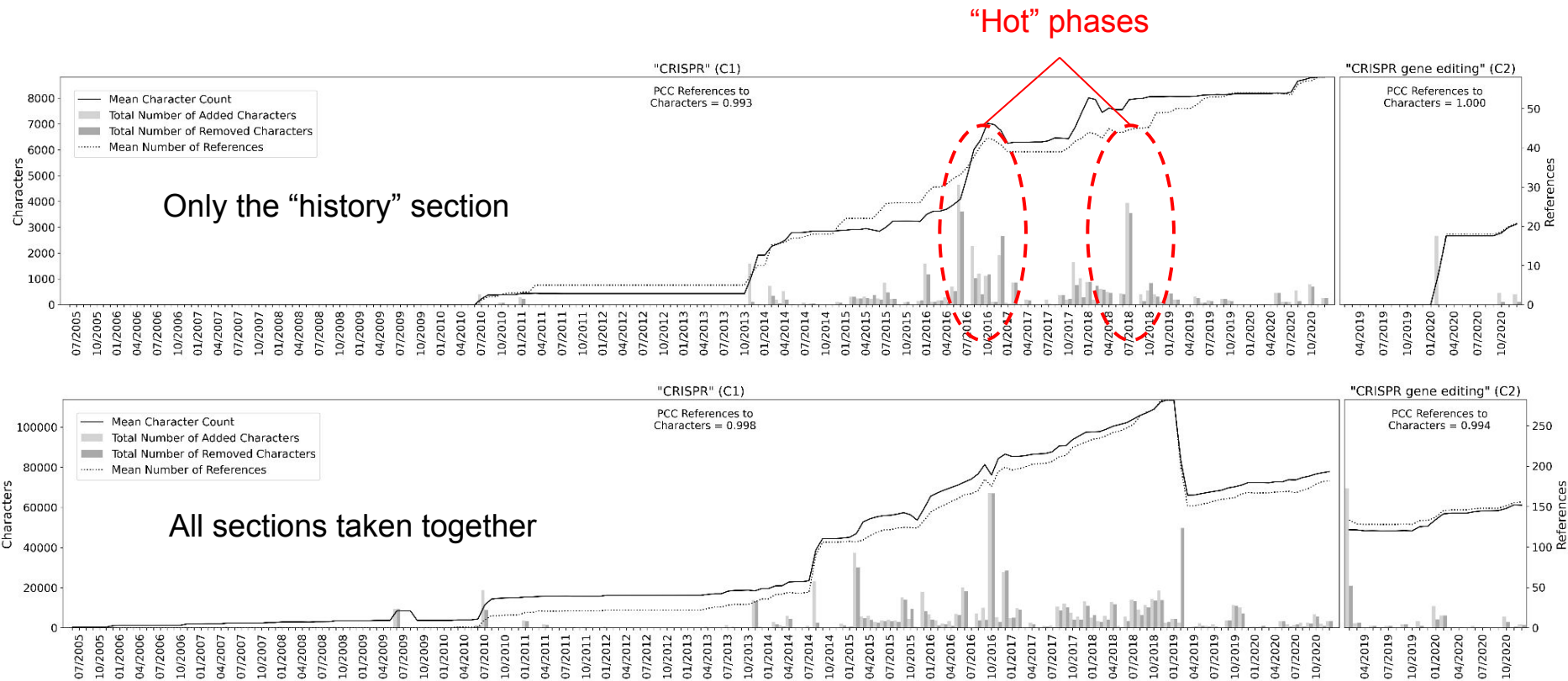*Y-axis (left): Number of character added/removed*

Only the "history" section

All sections taken together

Inspecting the editing landscape from far

Y-axis (right): Number of references

Only the "history" section

All sections taken together

Inspecting the editing landscape from far

"CRISPR" (C1) — PCC References to Characters = 0.993

Launch & growth of history section

Only the "history" section

"CRISPR gene editing" (C2) — PCC References to Characters = 1.000

"CRISPR" (C1) — PCC References to Characters = 0.998

All sections taken together

"CRISPR gene editing" (C2) — PCC References to Characters = 0.994

Legend: Mean Character Count; Total Number of Added Characters; Total Number of Removed Characters; Mean Number of References

Inspecting the editing landscape from far

"Hot" phases

Only the "history" section

All sections taken together

Inspecting the editing landscape from far

Plotting heroes from far

"CRISPR" (C1)

"CRISPR gene editing" (C2)

Plotting heroes from far
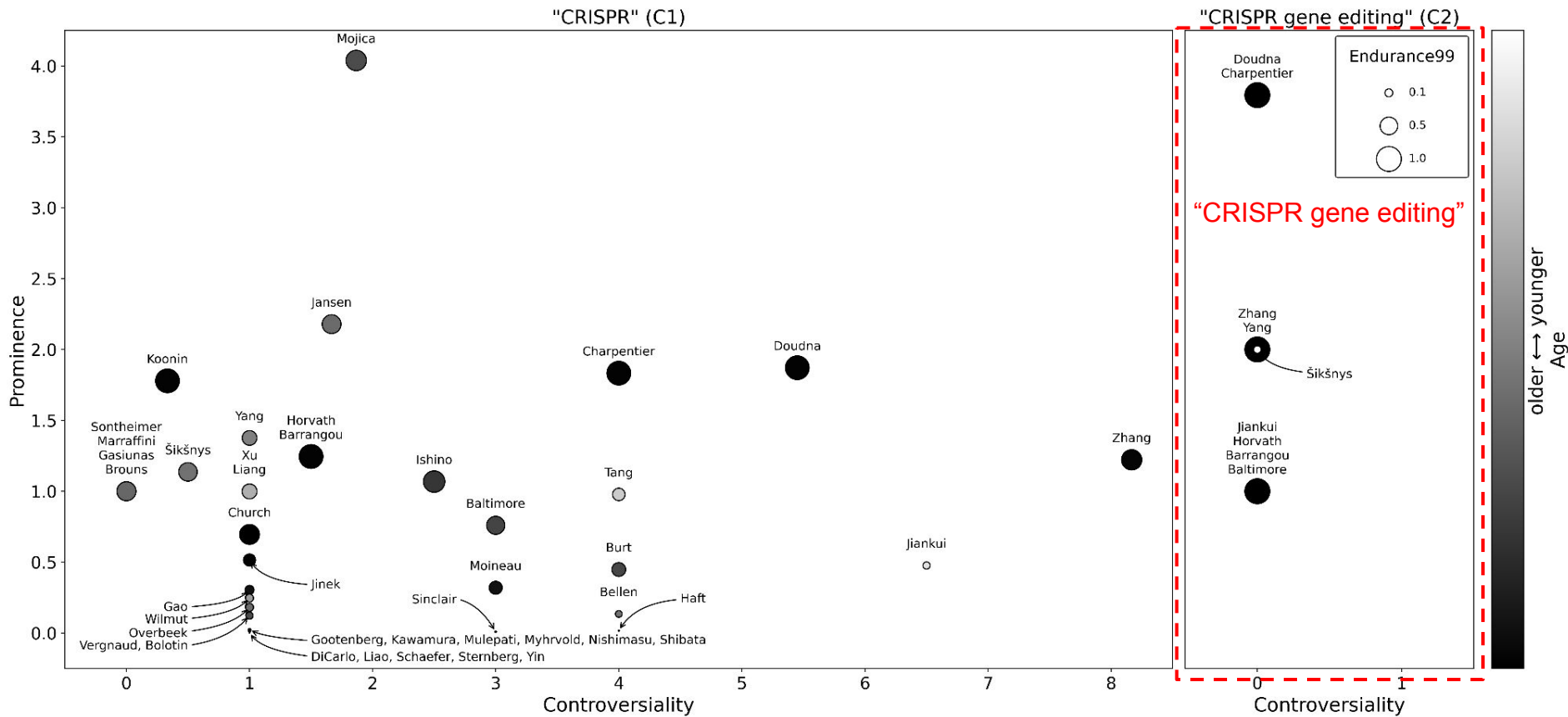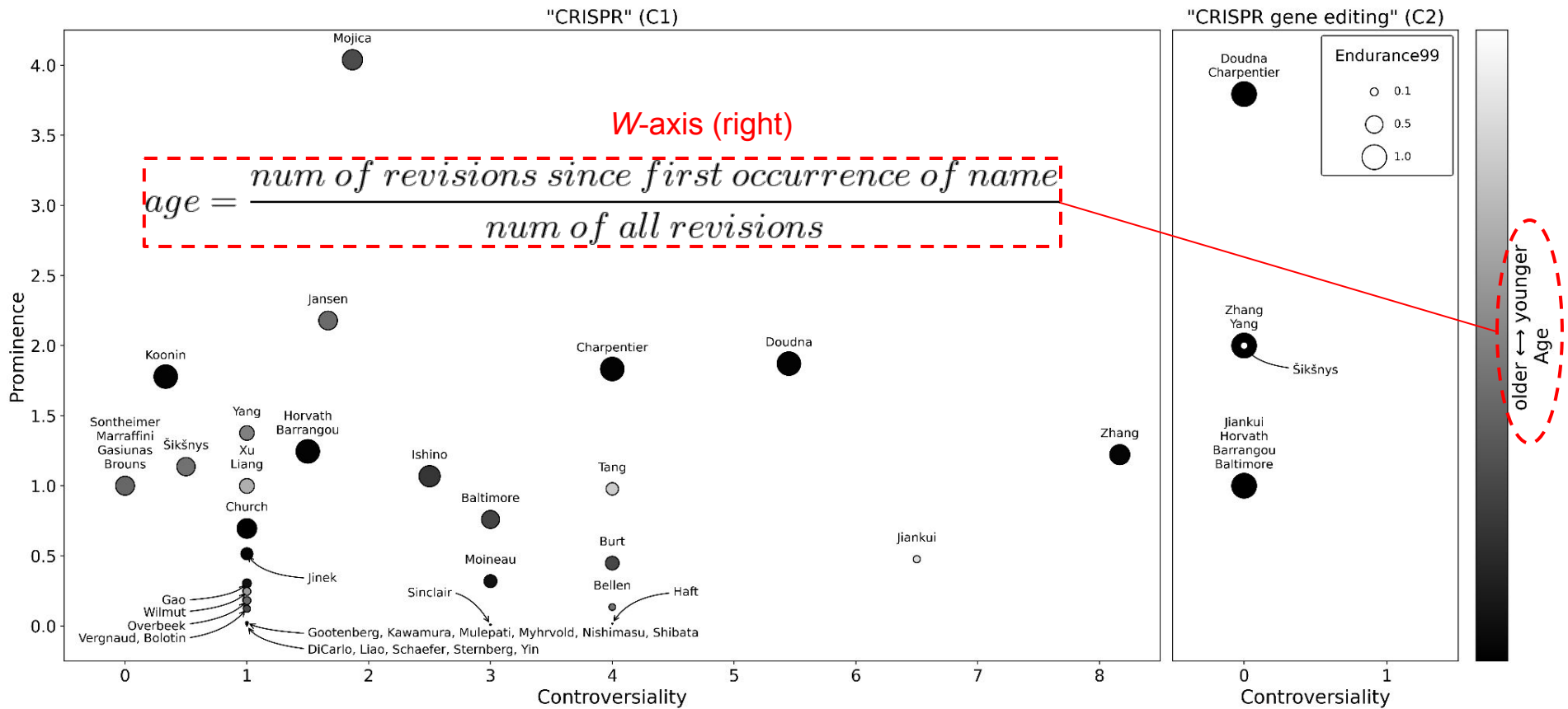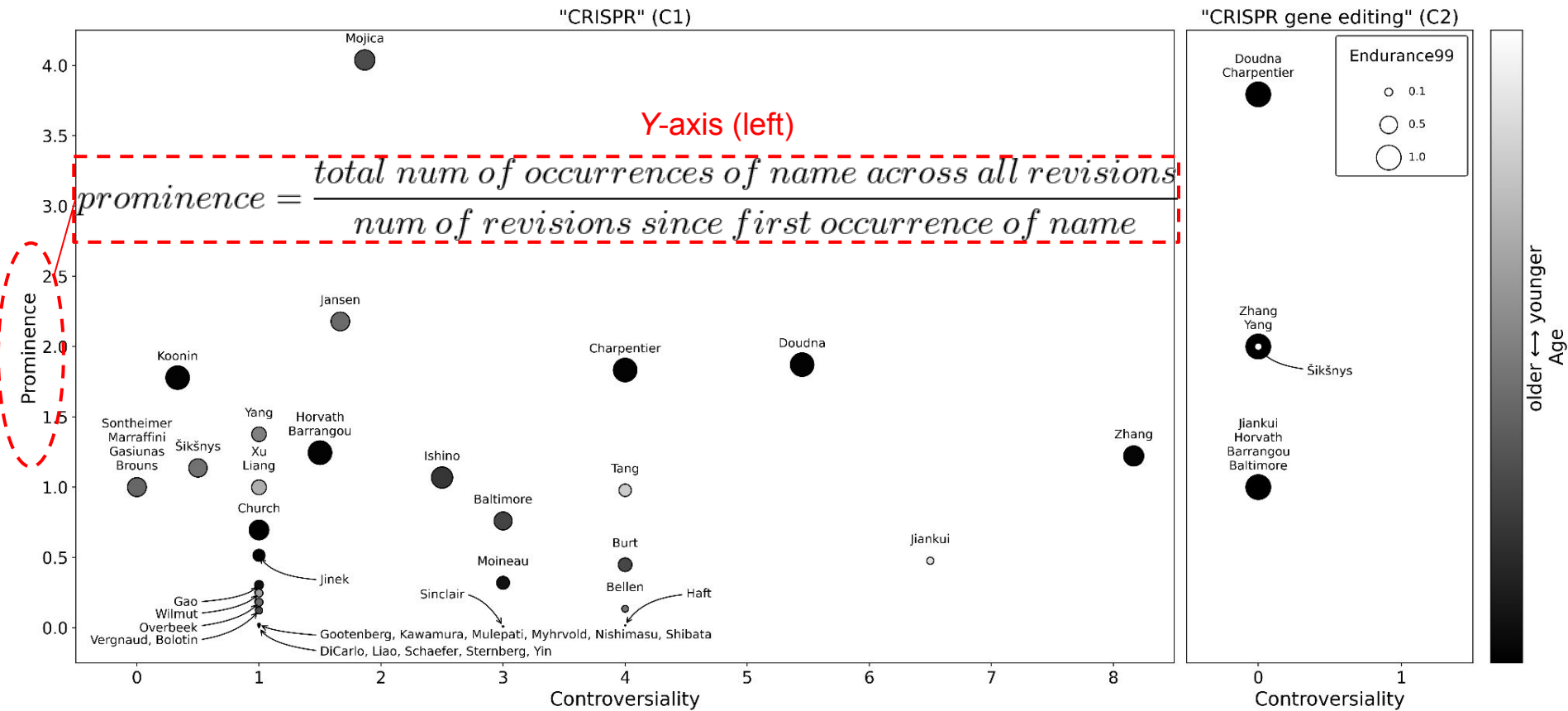
Plotting heroes from far

Plotting heroes from far

"CRISPR" (C1)

"CRISPR gene editing" (C2)

W-axis (right)

$$age = \frac{num\ of\ revisions\ since\ first\ occurrence\ of\ name}{num\ of\ all\ revisions}$$

Plotting heroes from far

"CRISPR" (C1)

Y-axis (left)

$$prominence = \frac{total\ num\ of\ occurrences\ of\ name\ across\ all\ revisions}{num\ of\ revisions\ since\ first\ occurrence\ of\ name}$$

"CRISPR gene editing" (C2)

Plotting heroes from far

$$controversiality = \frac{sum\ of\ neg\ change\ ratios\ in\ name\ freq\ from\ rev\ to\ rev}{-1}$$

X-axis

Plotting heroes from far

"CRISPR" (C1)

"CRISPR gene editing" (C2)

Z-axis (size)

$$endurance = \frac{num\ of\ revisions\ containing\ name}{num\ of\ all\ revisions}$$
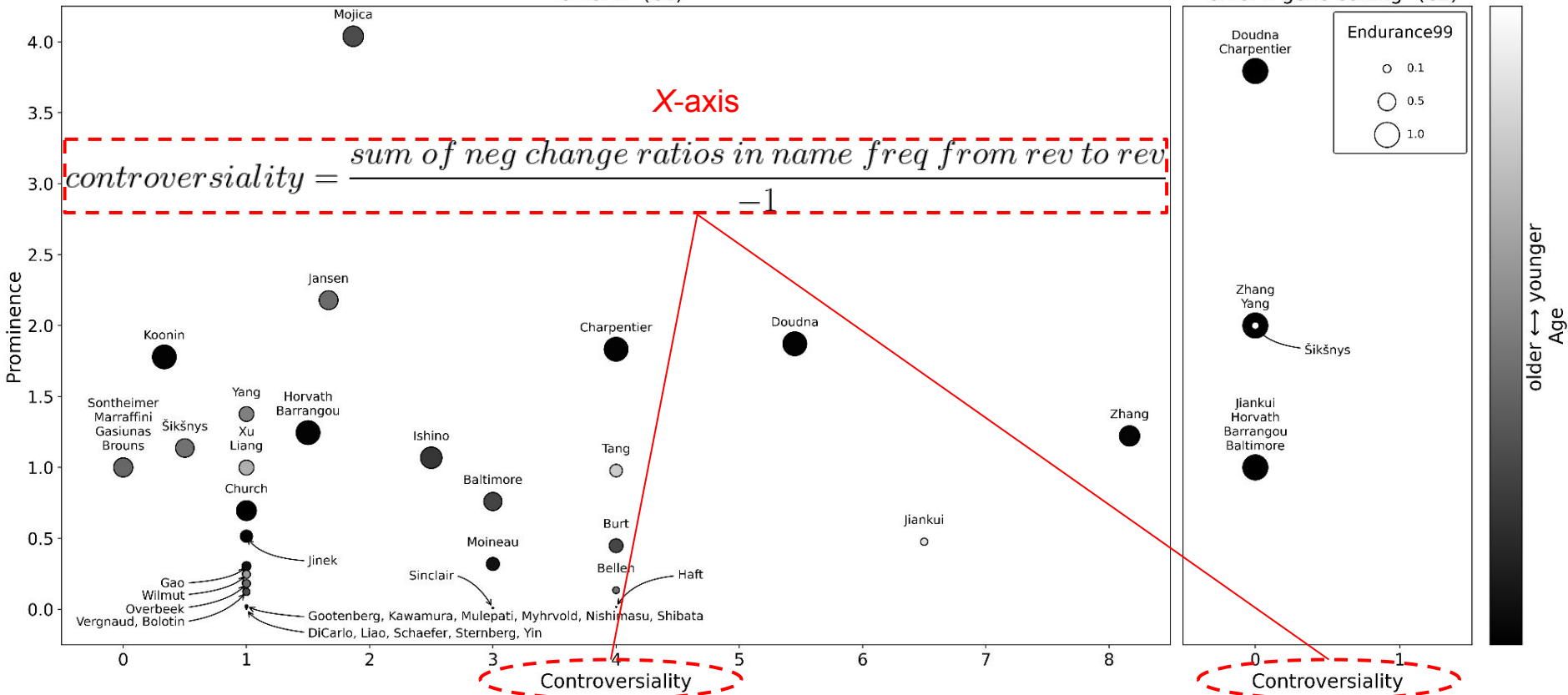
Plotting heroes from far

Plotting heroes from far

"CRISPR" (C1)

"CRISPR gene editing" (C2)

Less controversial
but quite prominent
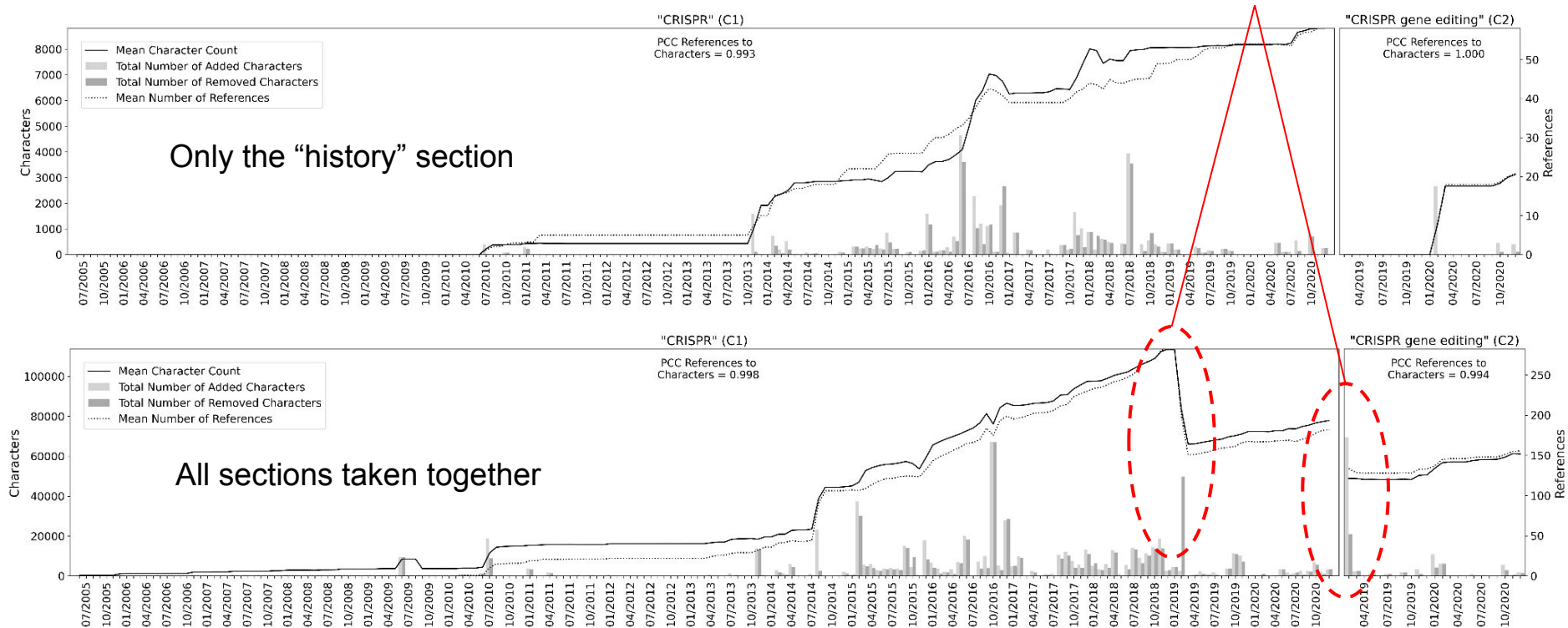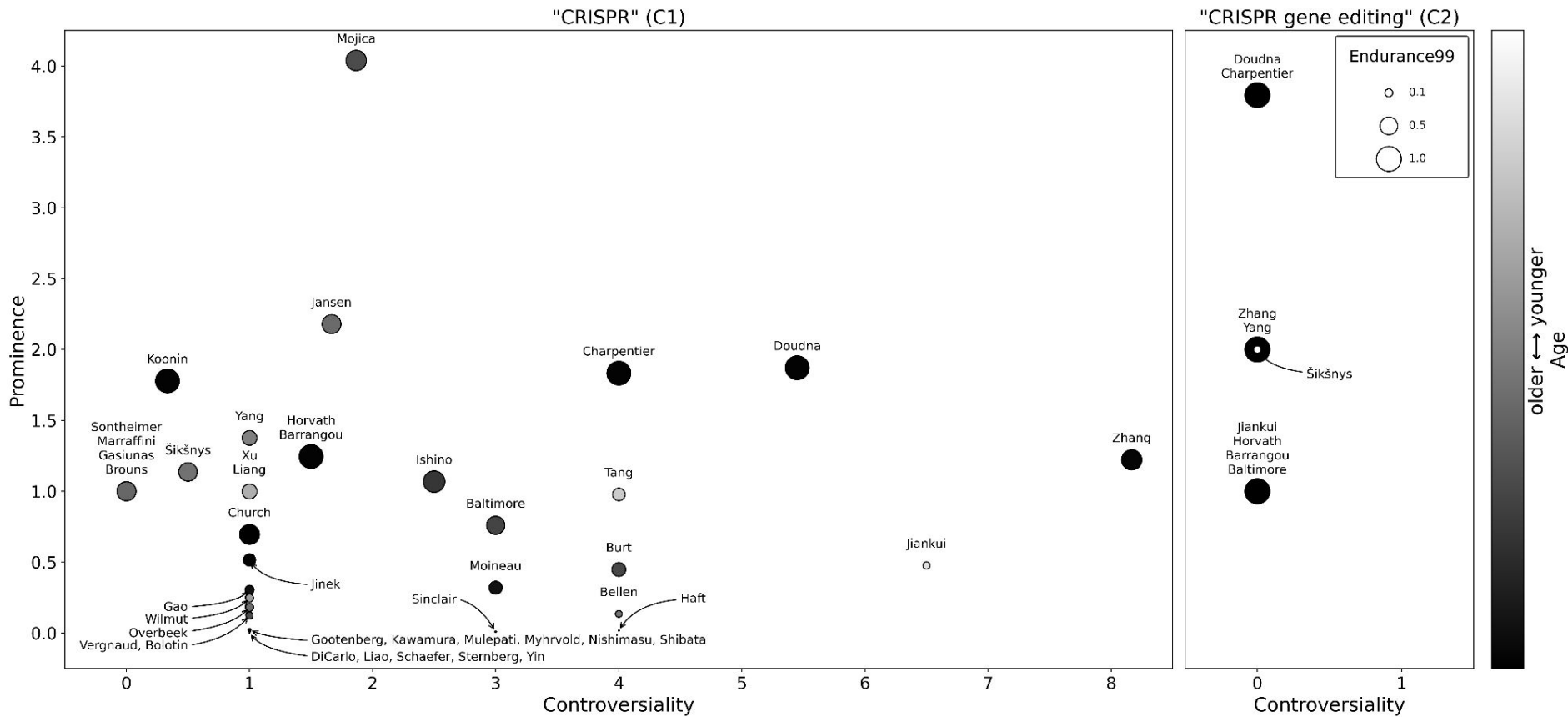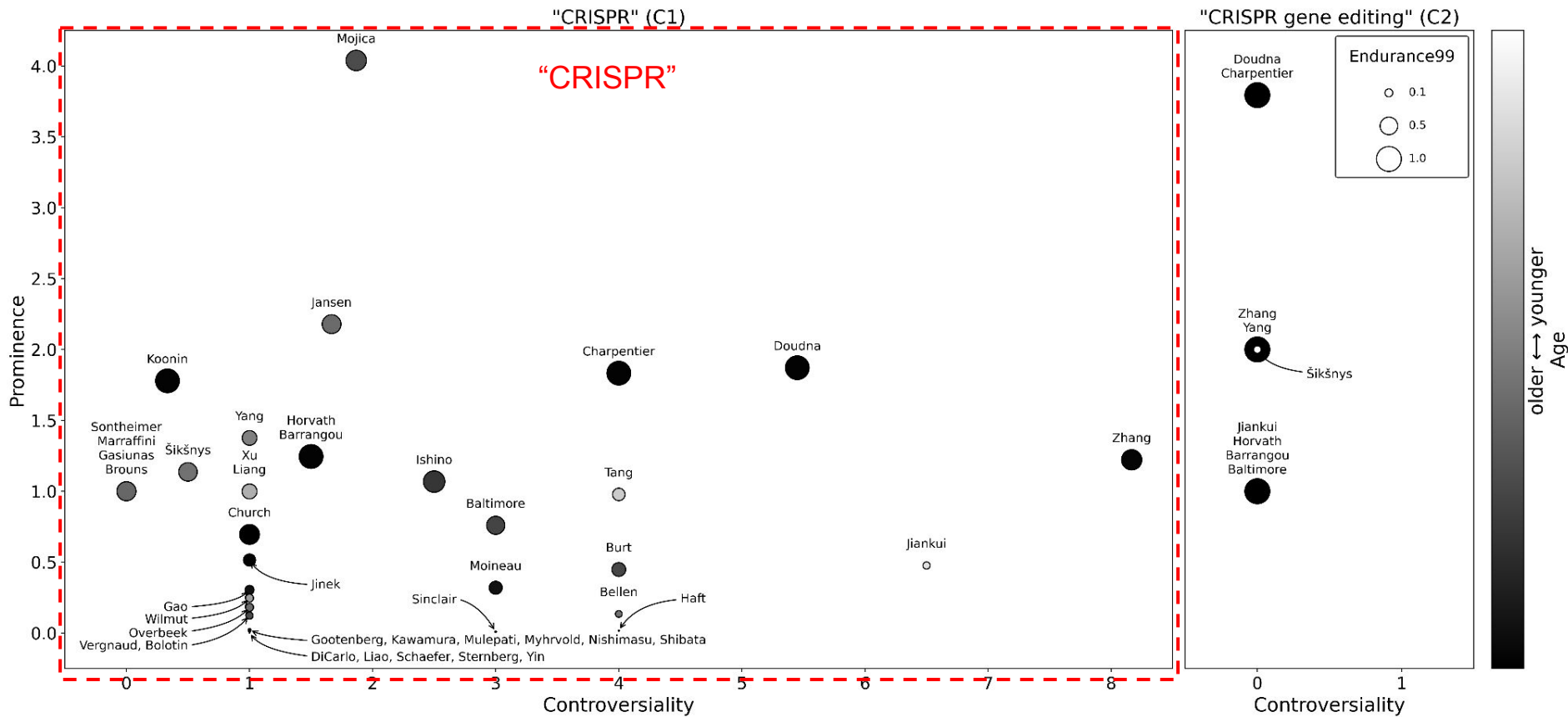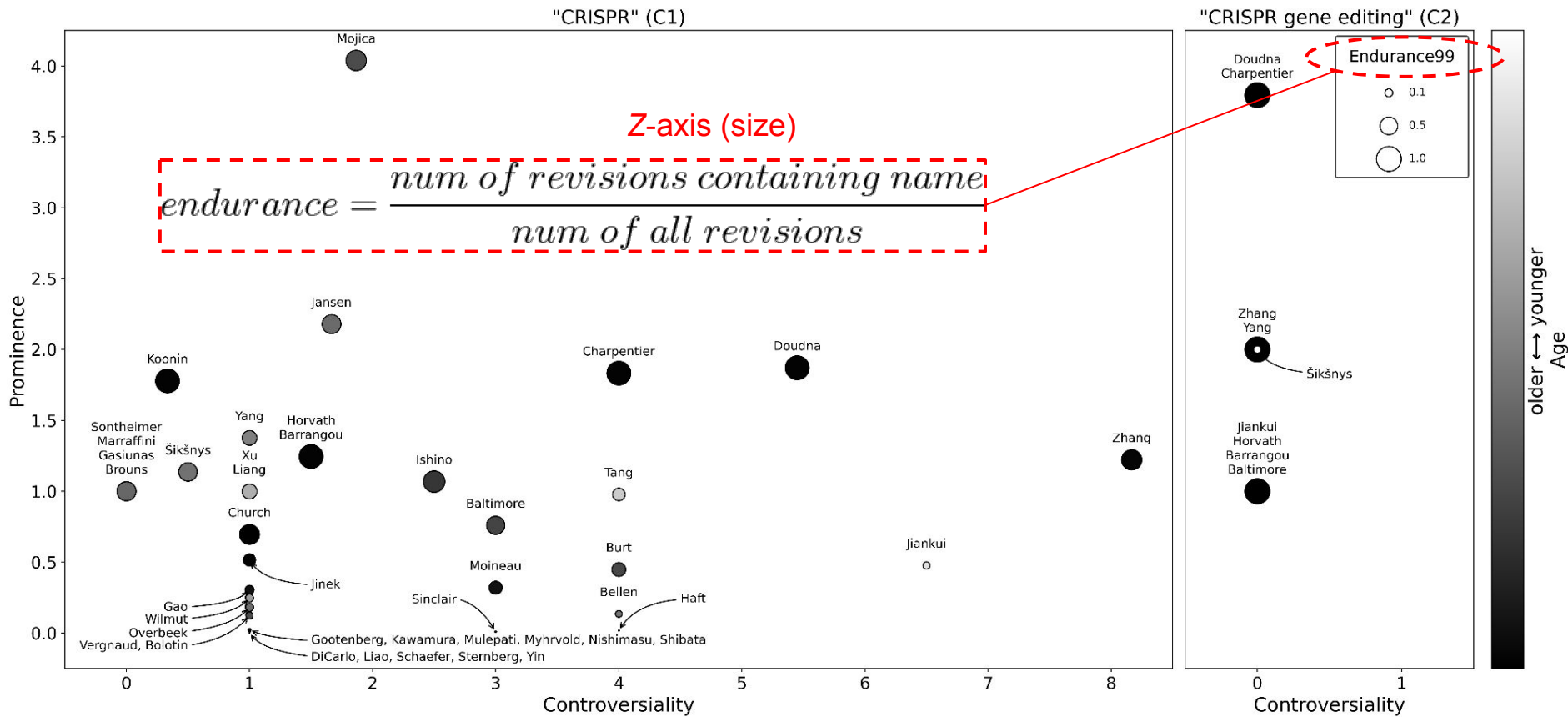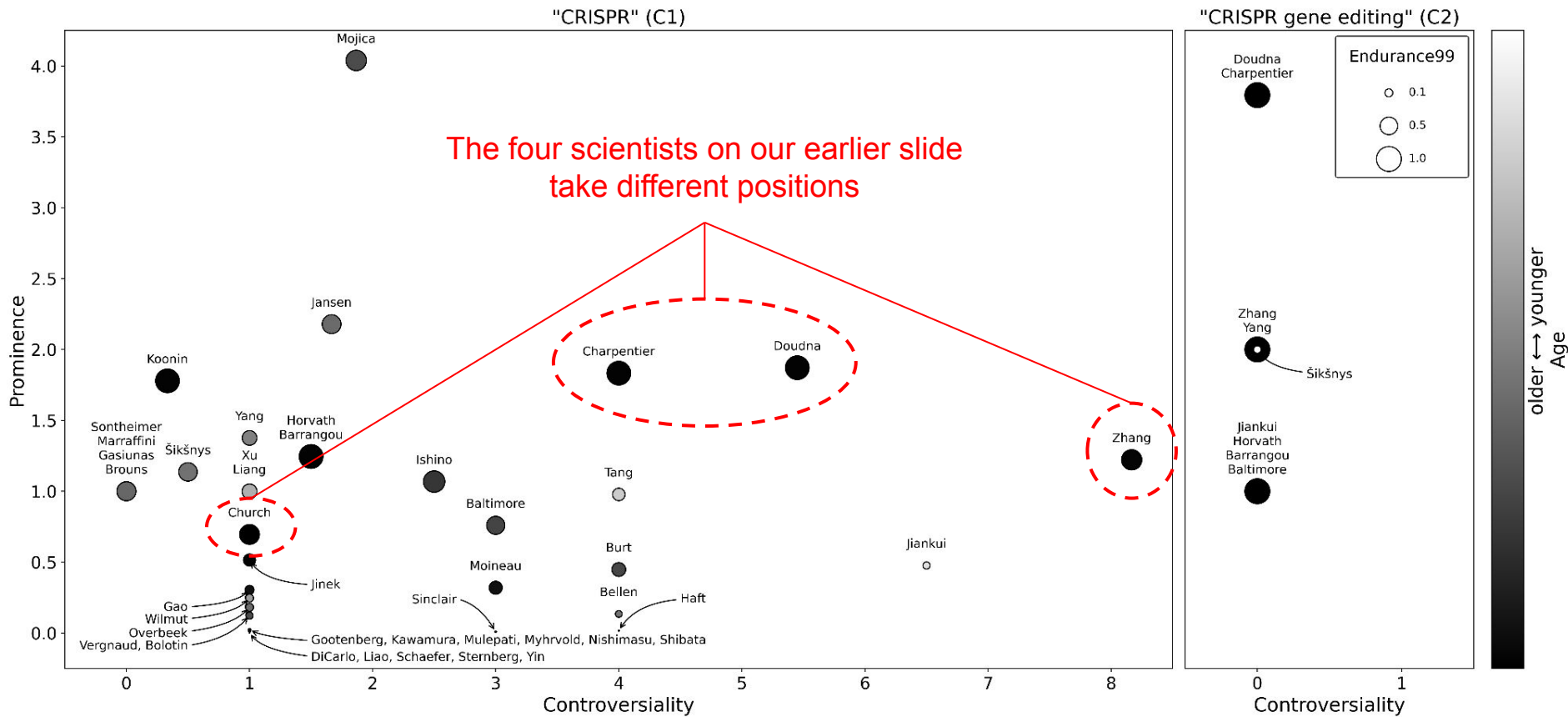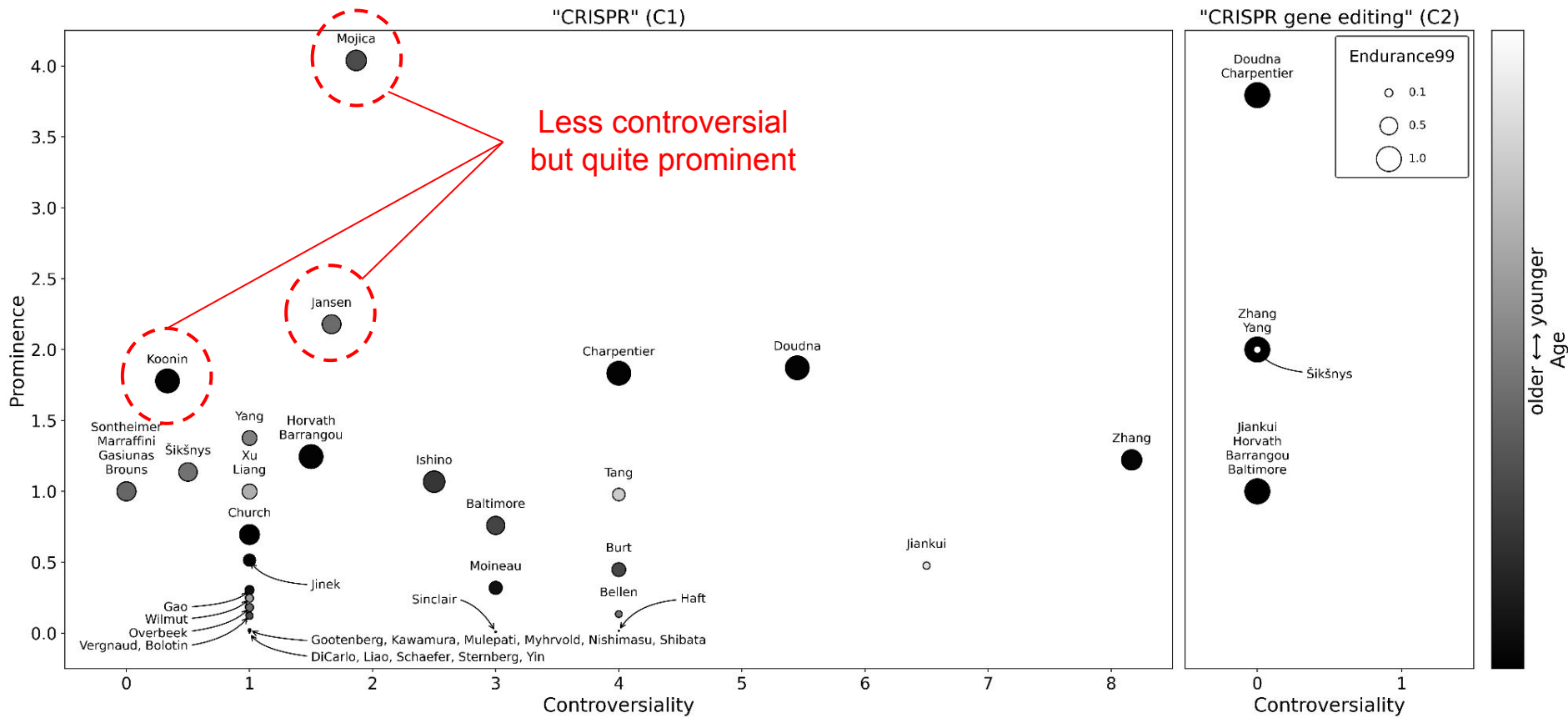
Plotting heroes from far
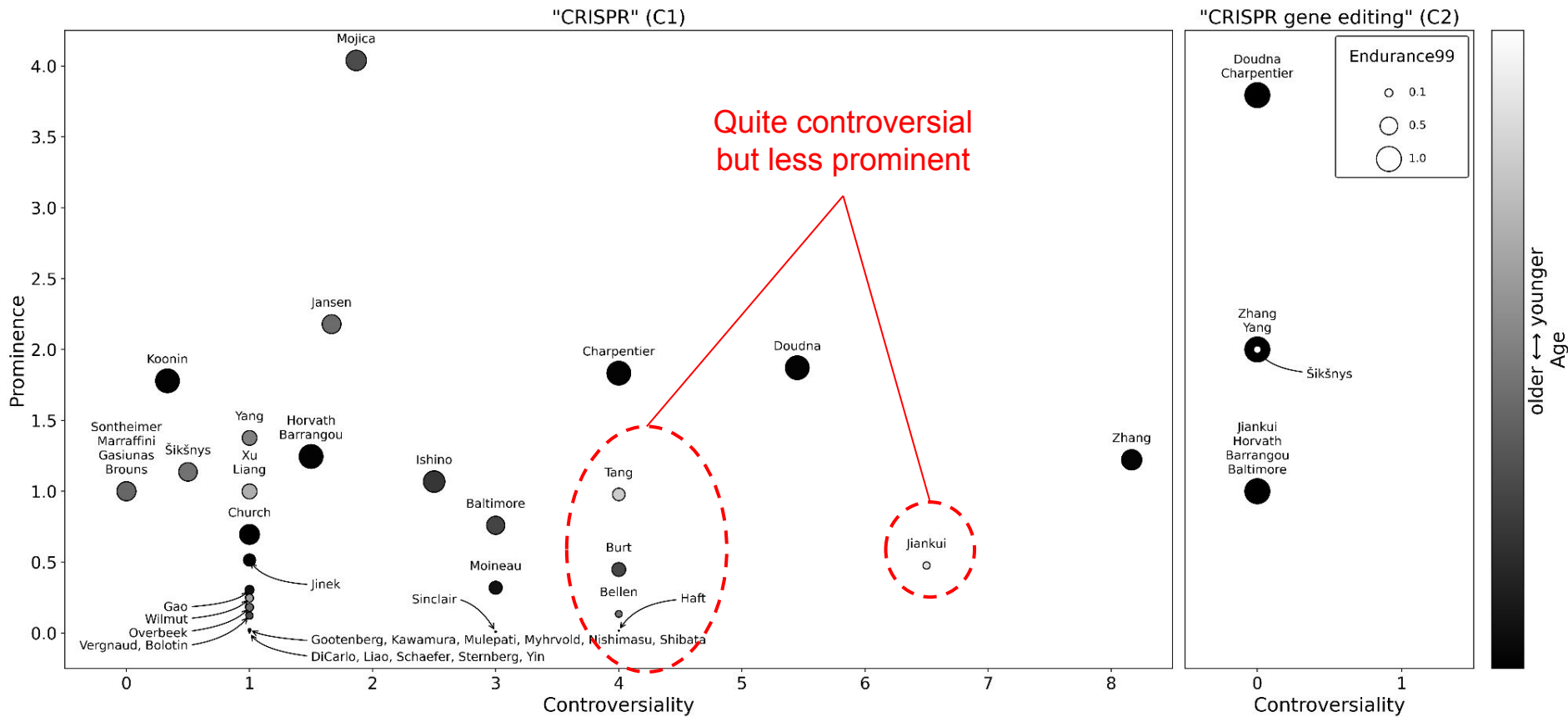
Reading from close

Three controversies

      Discovery of repeating DNA segments (CRISPR)

      Discovery that CRISPR is an immune defense system

      Development of CRISPR "genetic scissors"

Wikipedia editors negotiate their (different) accounts of "CRISPR Heroes"

      In and out of names and claims

Wikipedia's rules and policies used for (de-)legitimation of edits

      WP:MEDRS (Reliable sources)

      WP:PROMO (Self promotion)

Conclusions

Wikipedia not just a provider but also a *producer* of knowledge

    CRISPR entry/ revision = a "retrospective account" in its own right

Wikipedia works as a filter for relevance

    Claims, Names, Sources

Knowledge claims emerge from "socio-textual networking"

    Words, Reputation, Policies

<u>Adaptive</u> tool development worked well for us

    Flexibility to iterate between distant and close reading

Reflexivity: We also produced a retrospective account (of how Wikipedians account for CRISPR)

Thank you for listening!

Wikipedia Science Analytics

https://osf.io/ygur7

Arno Simons, Wolfgang Kircheis, Marion Schmidt, Martin Potthast, Benno Stein