The Role of Retention Time in Untargeted Metabolomics

Caparica-Lisbon, 21st September 2015

Jan Stanstrup†, Steffen Neumann‡, and Urška Vrhovšek†

†Department of Food Quality and Nutrition, Research and Innovation Centre, Fondazione Edmund Mach (FEM), San Michele all'Adige (TN), Italy

‡Department of Stress and Developmental Biology, Leibniz Institute of Plant Biochemistry, Halle, Germany

From feature to identified compound

Match to authentic standard

Orthogonal data: e.g. spectra + retention time

MSI LEVEL I

From feature to identified compound

Match to authentic standard

Orthogonal data: e.g. spectra + retention time

MSI LEVEL I

Match to public MS library

No orthogonal data: only spectra

MSI LEVEL II

From feature to identified compound

Match to authentic standard
Orthogonal data: e.g. spectra + retention time

MSI LEVEL I

Match to public MS library
No orthogonal data: only spectra

MSI LEVEL II

Compound not in any MS library

MSI LEVEL III/IV

From feature to identified compound

Match to public MS library

No orthogonal data: only spectra

MSI LEVEL II

MSI LEVEL II: match to public db

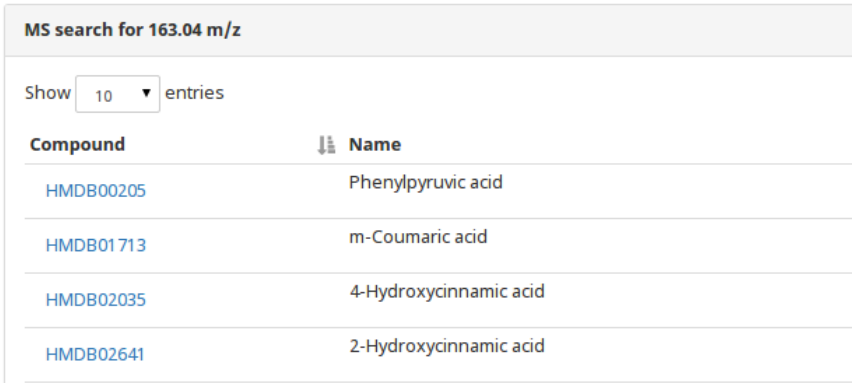
You got your match to a database spectra.

But are you sure that the spectra of your unknown couldn't match ***other*** isomers equally well?

MS search for 163.04 m/z

Show 10 entries

Compound	Name
HMDB00205	Phenylpyruvic acid
HMDB01713	m-Coumaric acid
HMDB02035	4-Hydroxycinnamic acid
HMDB02641	2-Hydroxycinnamic acid

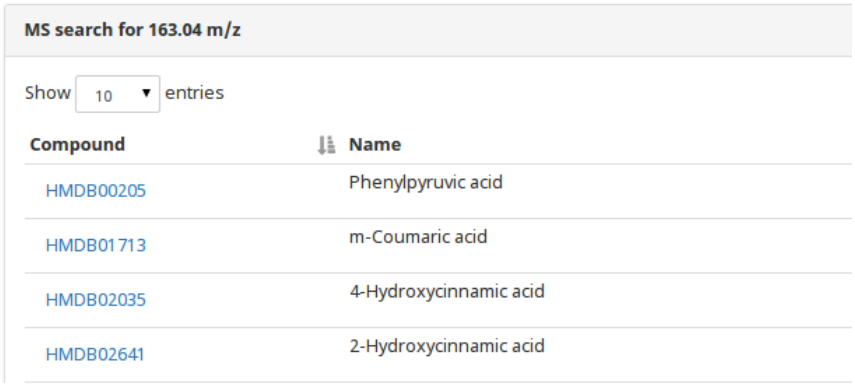


MS search for 163.04 m/z

Show 10 entries

Compound	Name
HMDB00205	Phenylpyruvic acid
HMDB01713	m-Coumaric acid
HMDB02035	4-Hydroxycinnamic acid
HMDB02641	2-Hydroxycinnamic acid

1) Examine mass spectra → No way to discriminate



MS search for 163.04 m/z

Show 10 entries

Compound	Name
HMDB00205	Phenylpyruvic acid
HMDB01713	m-Coumaric acid
HMDB02035	4-Hydroxycinnamic acid
HMDB02641	2-Hydroxycinnamic acid

1) Examine mass spectra → No way to discriminate
2) Check spectra libraries → Still cannot be sure

MS search for 163.04 m/z

Show 10 entries

Compound	Name
HMDB00205	Phenylpyruvic acid
HMDB01713	m-Coumaric acid
HMDB02035	4-Hydroxycinnamic acid
HMDB02641	2-Hydroxycinnamic acid

1) Examine mass spectra → No way to discriminate
2) Check spectra libraries → Still cannot be sure

"Piled Higher and Deeper" by Jorge Cham www.phdcomics.com

MS search for 163.04 m/z

Show 10 entries

Compound	Name
HMDB00205	Phenylpyruvic acid
HMDB01713	m-Coumaric acid
HMDB02035	4-Hydroxycinnamic acid
HMDB02641	2-Hydroxycinnamic acid

1) Examine mass spectra → No way to discriminate
2) Check spectra libraries → Still cannot be sure
3) Buy all standards (if available) → Expensive!

"Piled Higher and Deeper" by Jorge Cham www.phdcomics.com

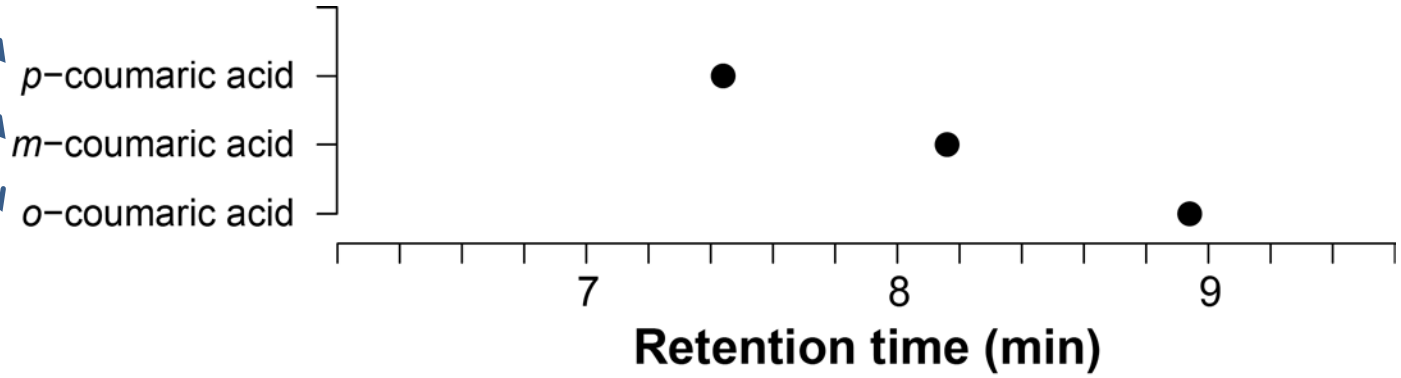
MS search for 163.04 m/z

Show 10 entries

Compound	Name
HMDB00205	Phenylpyruvic acid
HMDB01713	m-Coumaric acid
HMDB02035	4-Hydroxycinnamic acid
HMDB02641	2-Hydroxycinnamic acid

1) Examine mass spectra → No way to discriminate
2) Check spectra libraries → Still cannot be sure
3) Buy all standards (if available) → Expensive!

"Piled Higher and Deeper" by Jorge Cham www.phdcomics.com



MS search for 163.04 m/z

Show 10 entries

Compound	Name
HMDB00205	Phenylpyruvic acid
HMDB01713	m-Coumaric acid
HMDB02035	4-Hydroxycinnamic acid
HMDB02641	2-Hydroxycinnamic acid

1) Examine mass spectra → No way to discriminate
2) Check spectra libraries → Still cannot be sure
3) Buy all standards (if available) → Expensive!

"Piled Higher and Deeper" by Jorge Cham www.phdcomics.com

So can we at least limit the number of isomers by using the information we already have?

Why retention time information is usually neglected

Retention time is specific to the chromatographic system and there are no RT references

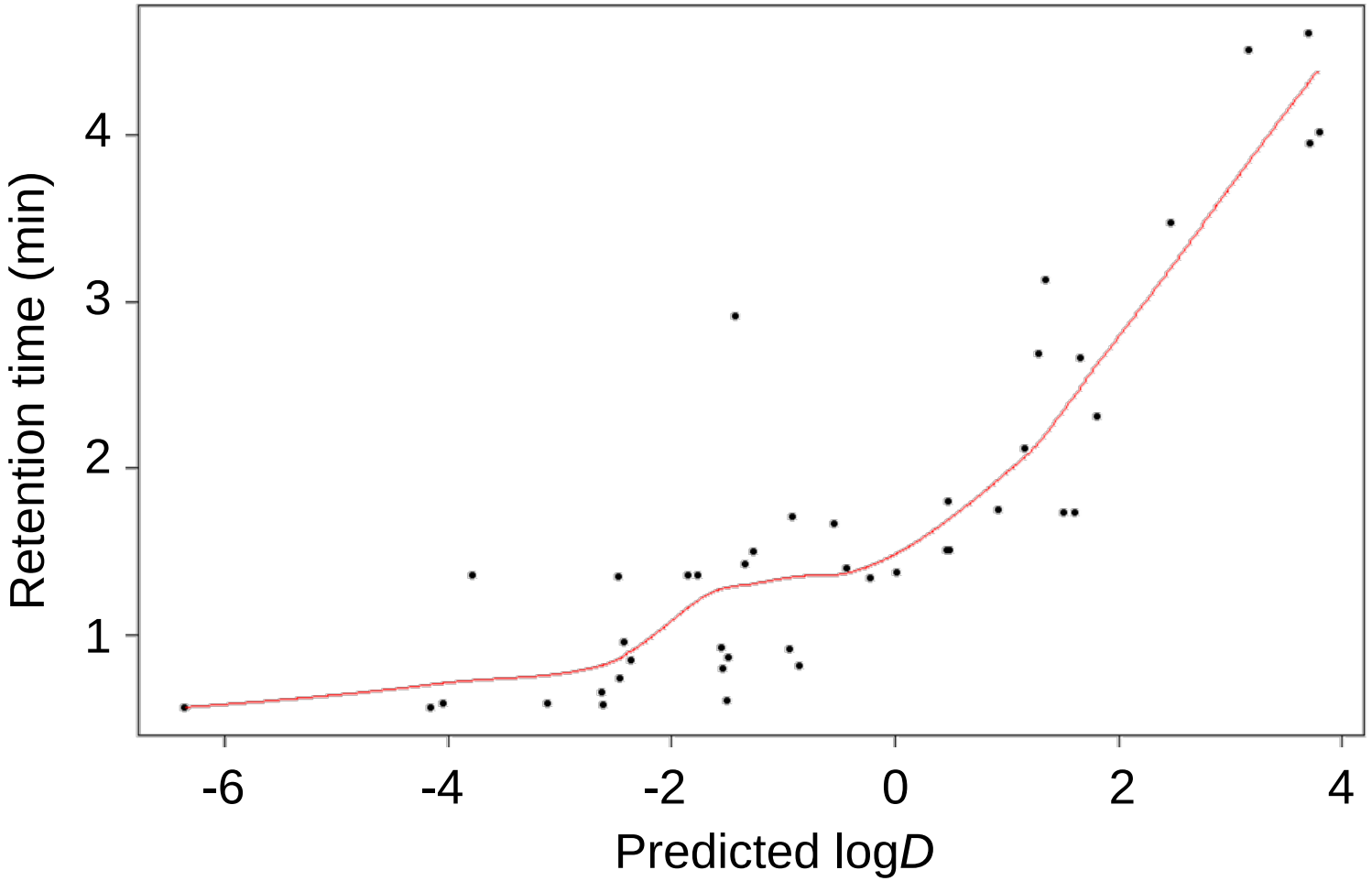
No coordinated efforts to share and exploit RT info.

Current approaches to retention time prediction and their characteristics

	Prediction from structure (QSRR)	**Projection from isocratic runs**[a]
Accuracy	Low	Very high
Universality (any structure)	High	Low
Additional lab work	None	Lots
Model complexity	High	Low

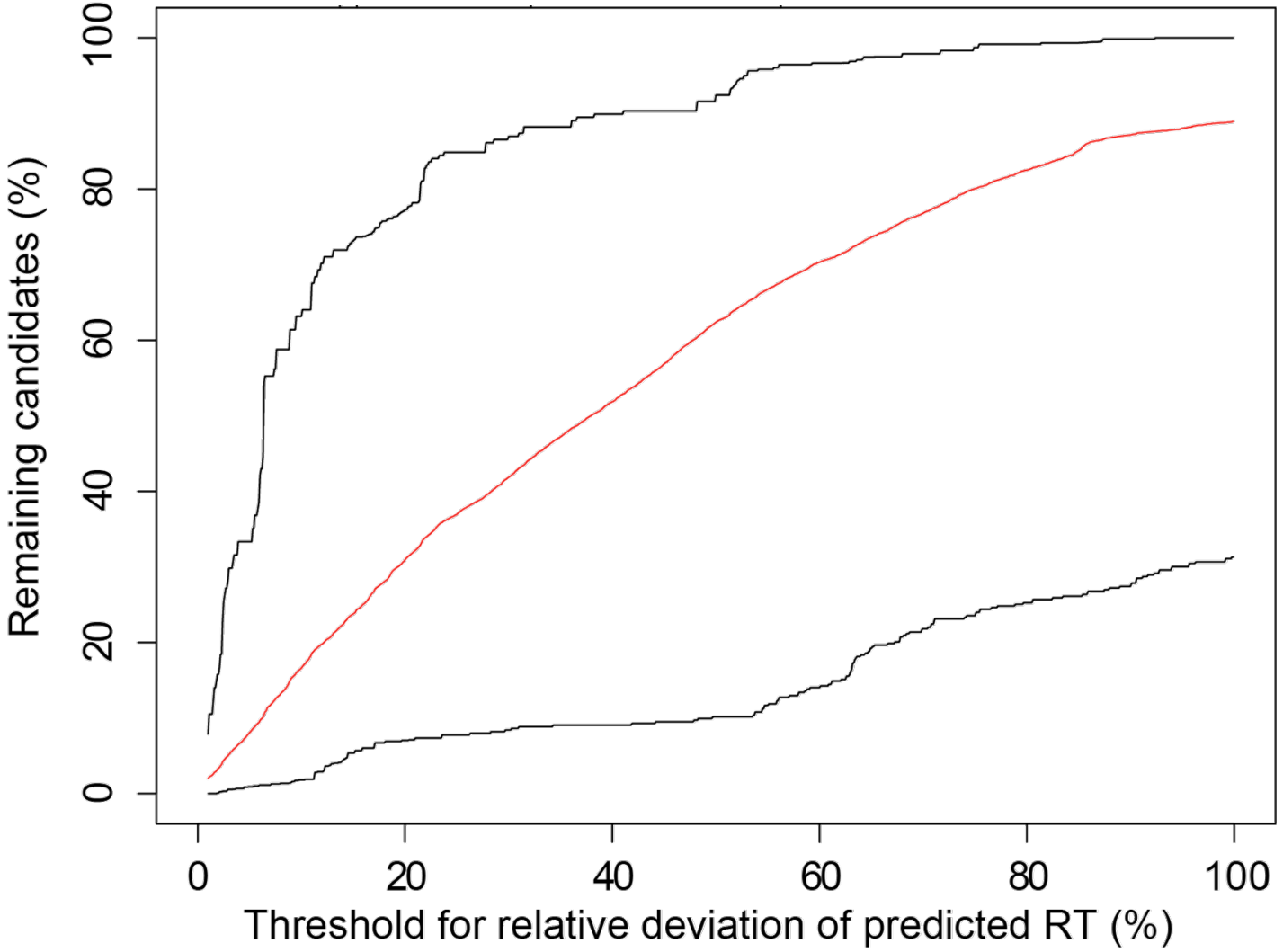
[a]Abate-Pella D, et al. (2015). Retention Projection Enables Accurate Calculation of Liquid Chromatographic Retention Times Across Labs and Methods. *J Chromatogr.*

Simplest QSRR model

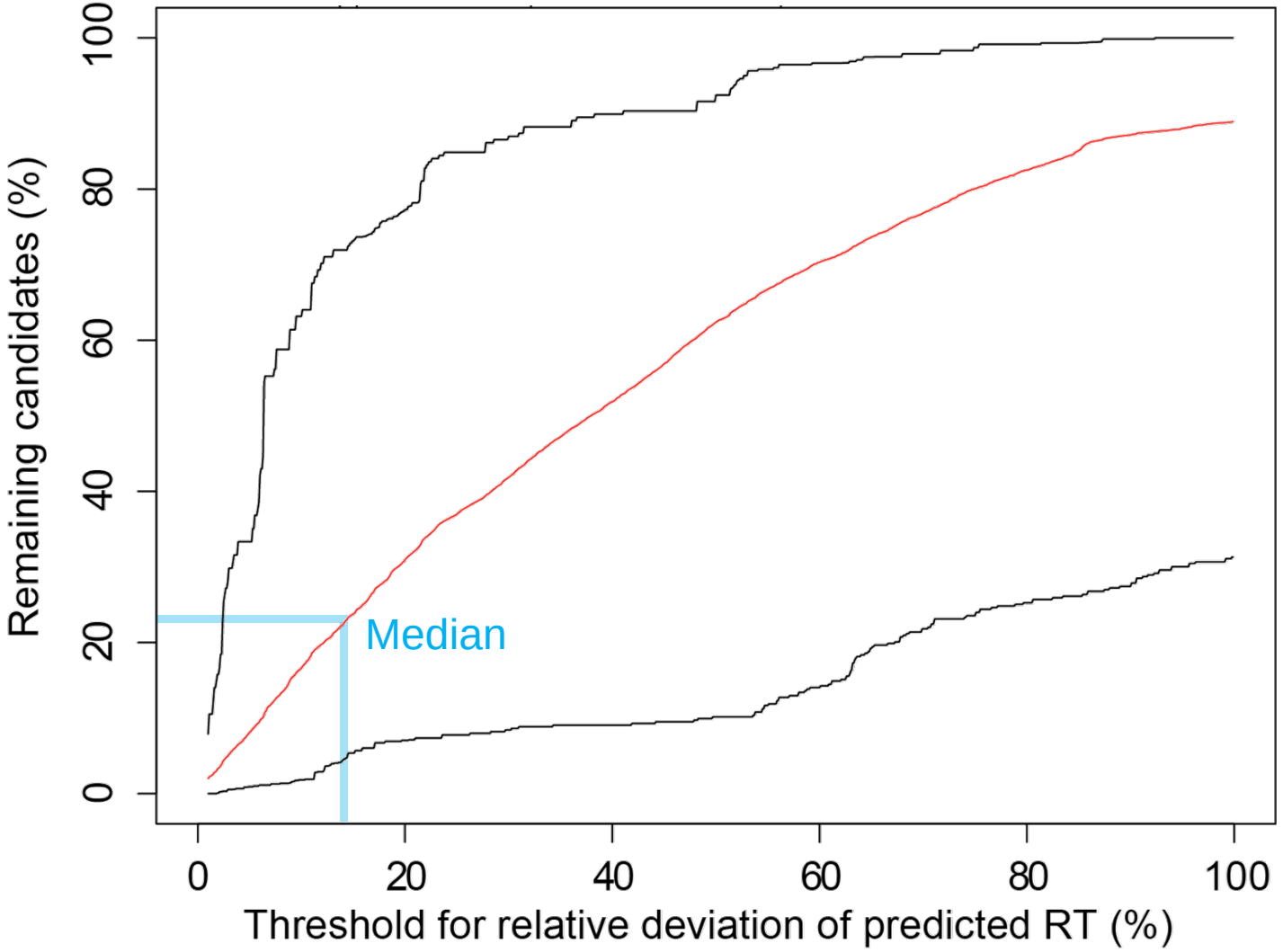


Stanstrup, J., Gerlich, M., Dragsted, L. O., & Neumann, S. (2013). Metabolite profiling and beyond: Approaches for the rapid processing and annotation of human blood serum mass spectrometry data Metabolomics and Metabolite Profiling. *Analytical and Bioanalytical Chemistry*, *405*(15), 5037–5048.

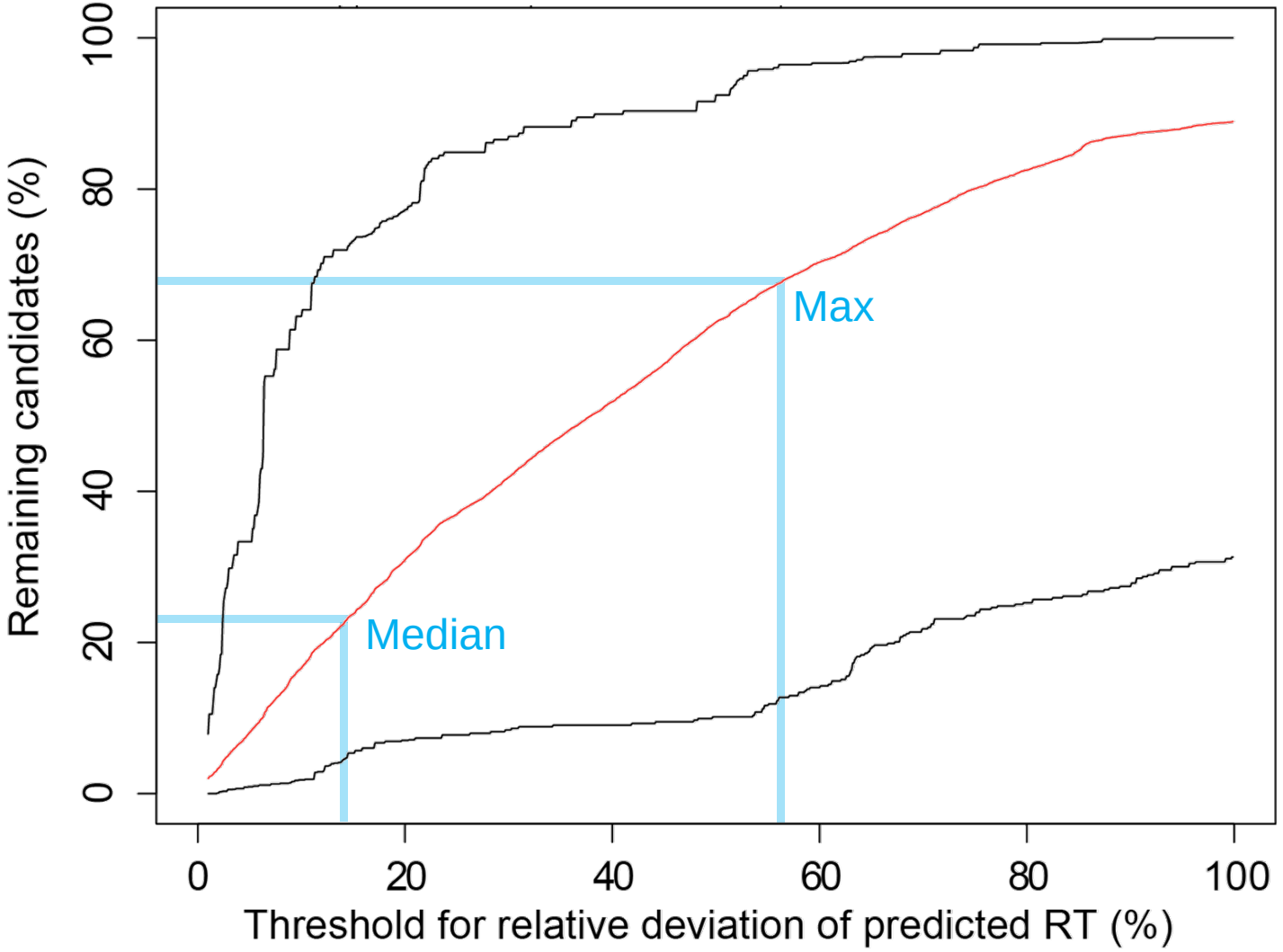
Candidates after filtering as a function of prediction accuracy



Candidates after filtering as a function of prediction accuracy



Candidates after filtering as a function of prediction accuracy



FONDAZIONE EDMUND MACH

THE VELUX FOUNDATIONS
VILLUM FONDEN ✕ VELUX FONDEN

Previous approaches to retention time prediction

	Prediction from structure (QSRR)	**Projection from isocratic runs**[a]
Accuracy	Low	Very high
Universality (any structure)	High	Low
Additional lab work	None	Lots
Model complexity	High	Low

[a]Abate-Pella D, et al. (2015). Retention Projection Enables Accurate Calculation of Liquid Chromatographic Retention Times Across Labs and Methods. *J Chromatogr.*

FONDAZIONE EDMUND MACH

THE VELUX FOUNDATIONS
VILLUM FONDEN × VELUX FONDEN

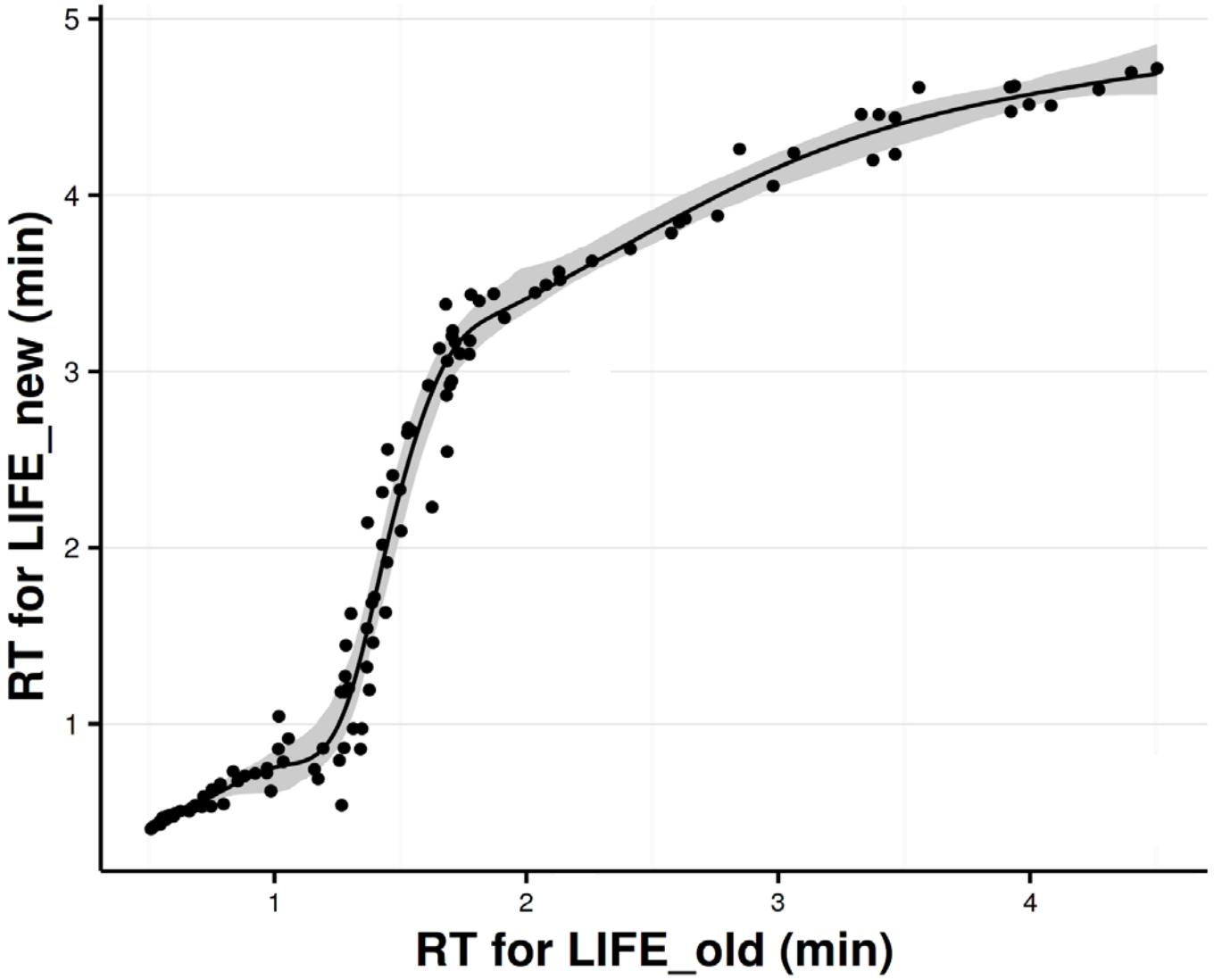
Previous approaches to retention time prediction

	Prediction from structure (QSRR)	**Projection from isocratic runs**[a]	**PredRet**[b]
Accuracy	Low	Very high	Medium to high
Universality (any structure)	High	Low	Low to medium
Additional lab work	None	Lots	None
Model complexity	High	Low	Low

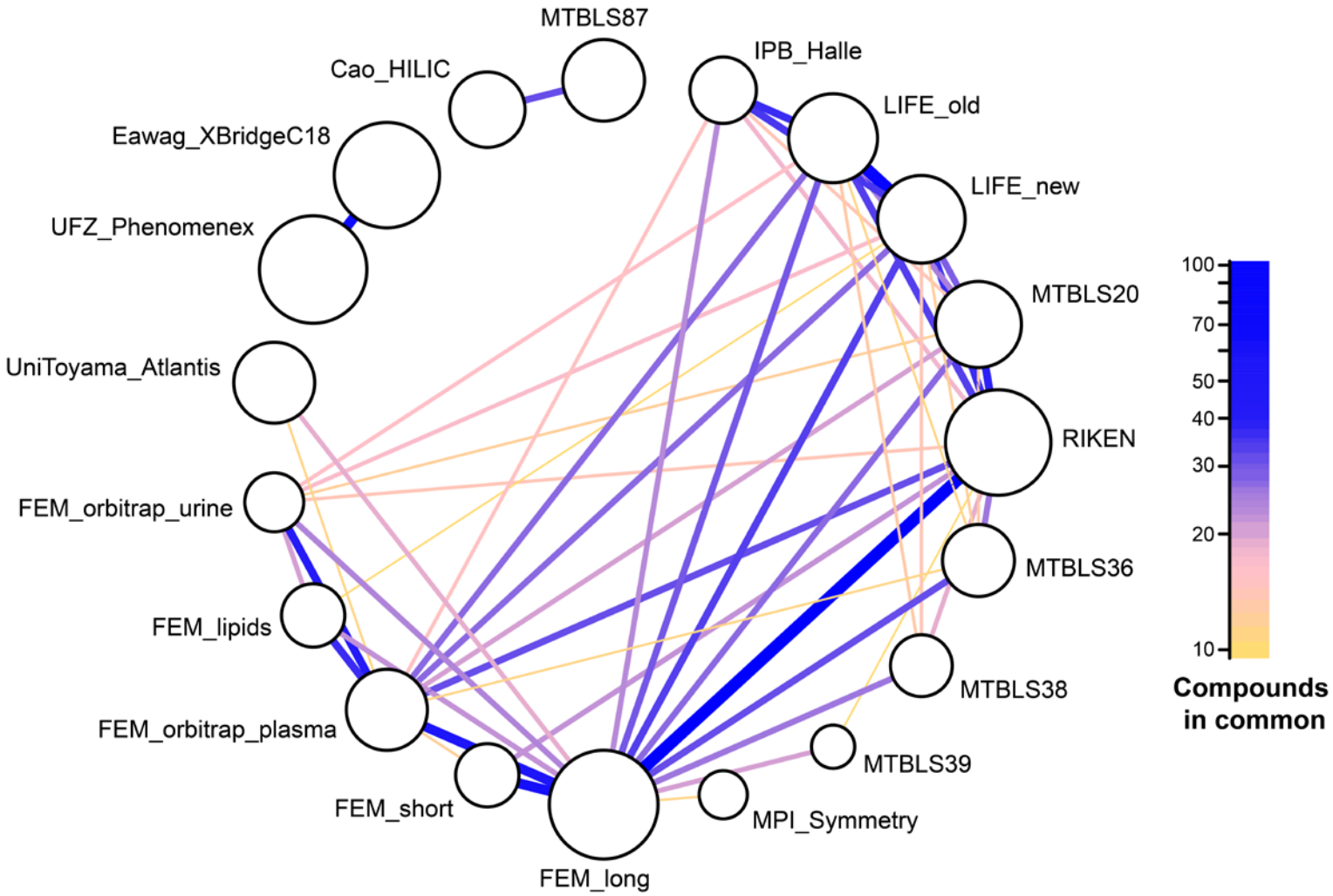
[a]Abate-Pella D, et al. (2015). Retention Projection Enables Accurate Calculation of Liquid Chromatographic Retention Times Across Labs and Methods. *J Chromatogr.*

[b]Stanstrup, J., Neumann, S., & Vrhovsek, U. (2015). PredRet: Prediction of Retention Time by Direct Mapping between Multiple Chromatographic Systems. *Analytical Chemistry*.

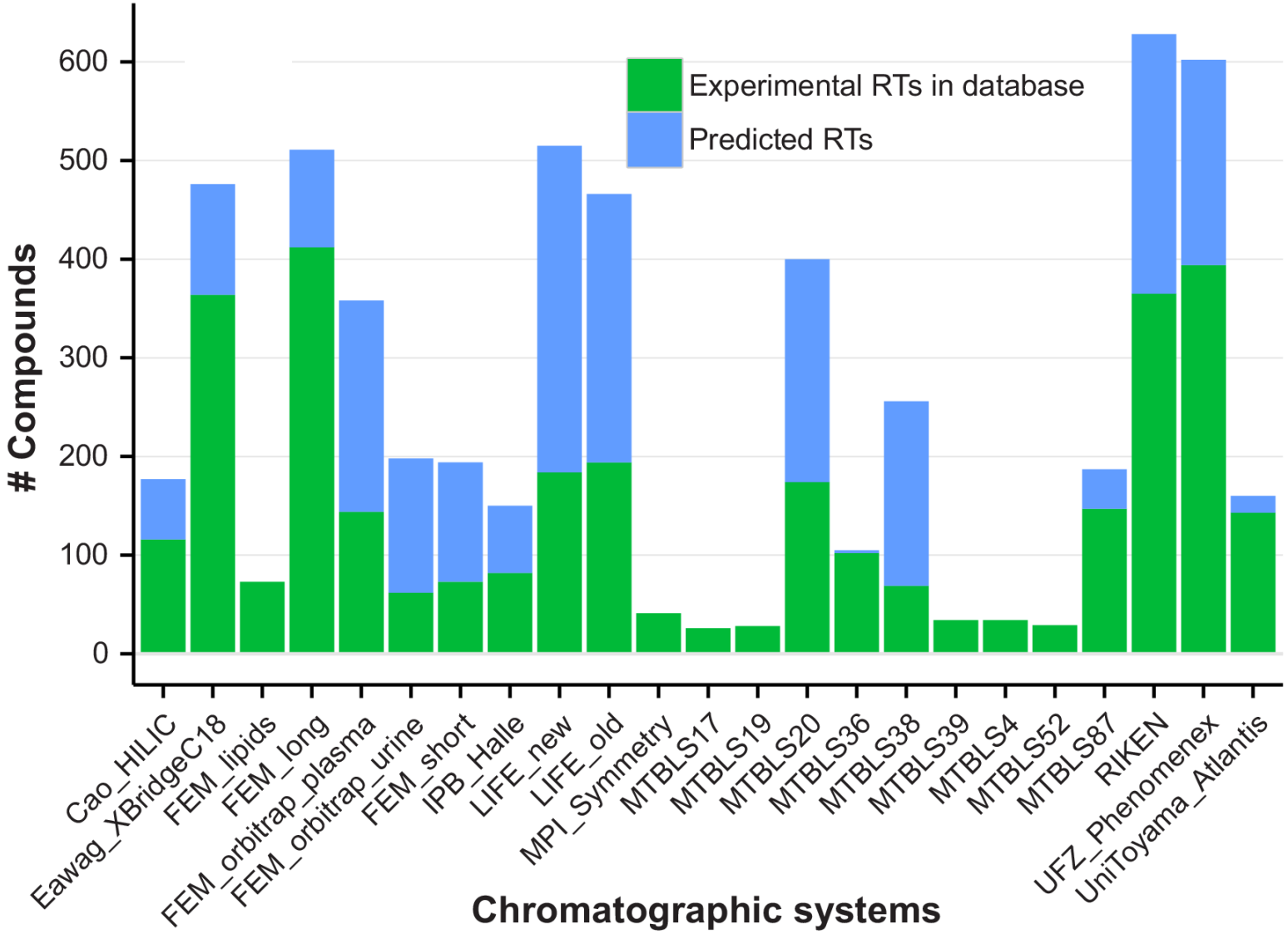
PredRet models



Compound overlap between systems in PredRet



Number of RTs in database and predictions made



MS search for 163.04 m/z

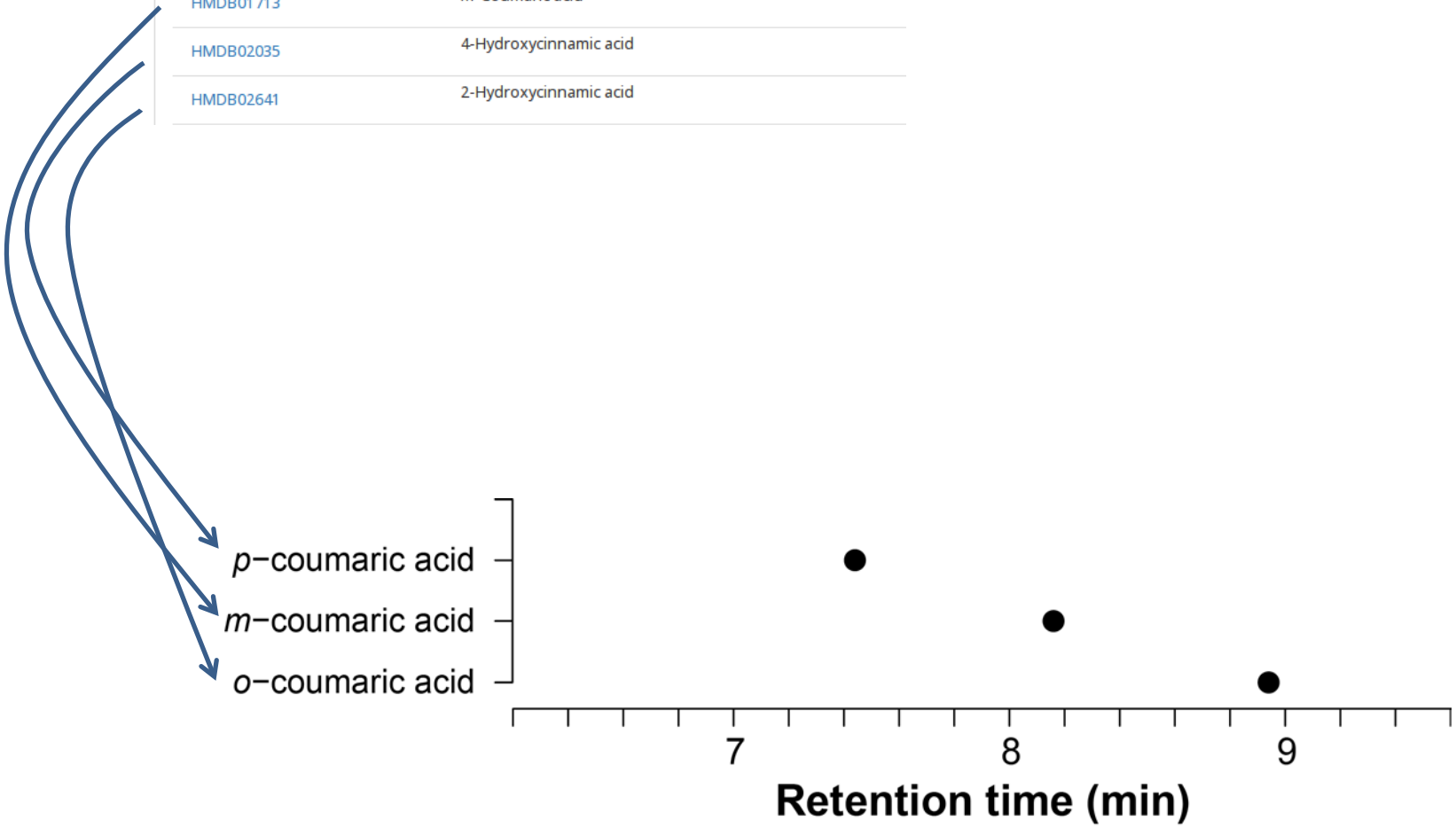
Show 10 entries

Compound	Name
HMDB00205	Phenylpyruvic acid
HMDB01713	m-Coumaric acid
HMDB02035	4-Hydroxycinnamic acid
HMDB02641	2-Hydroxycinnamic acid

MS search for 163.04 m/z

Show 10 entries

Compound	Name
HMDB00205	Phenylpyruvic acid
HMDB01713	m-Coumaric acid
HMDB02035	4-Hydroxycinnamic acid
HMDB02641	2-Hydroxycinnamic acid



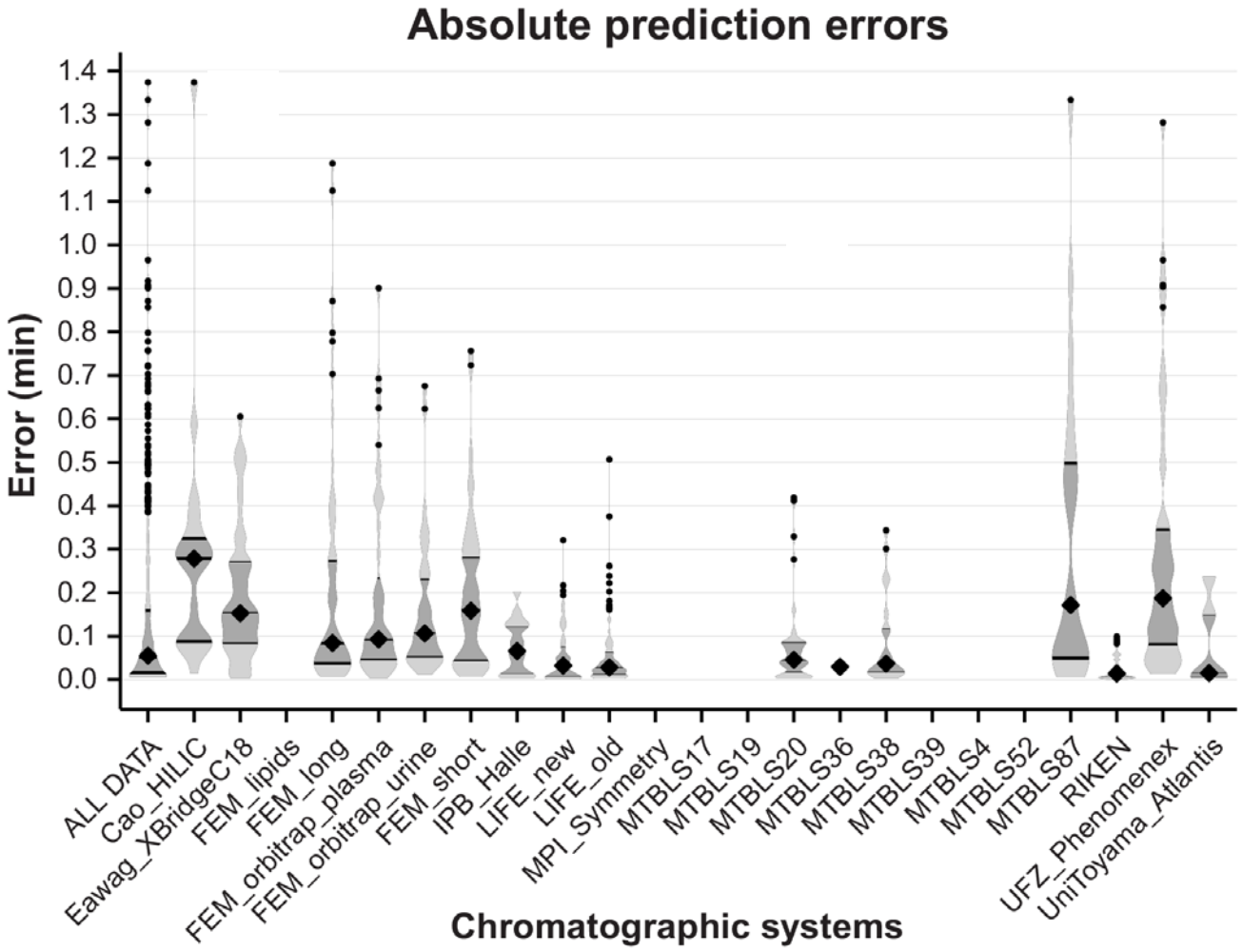
MS search for 163.04 m/z

Show 10 entries

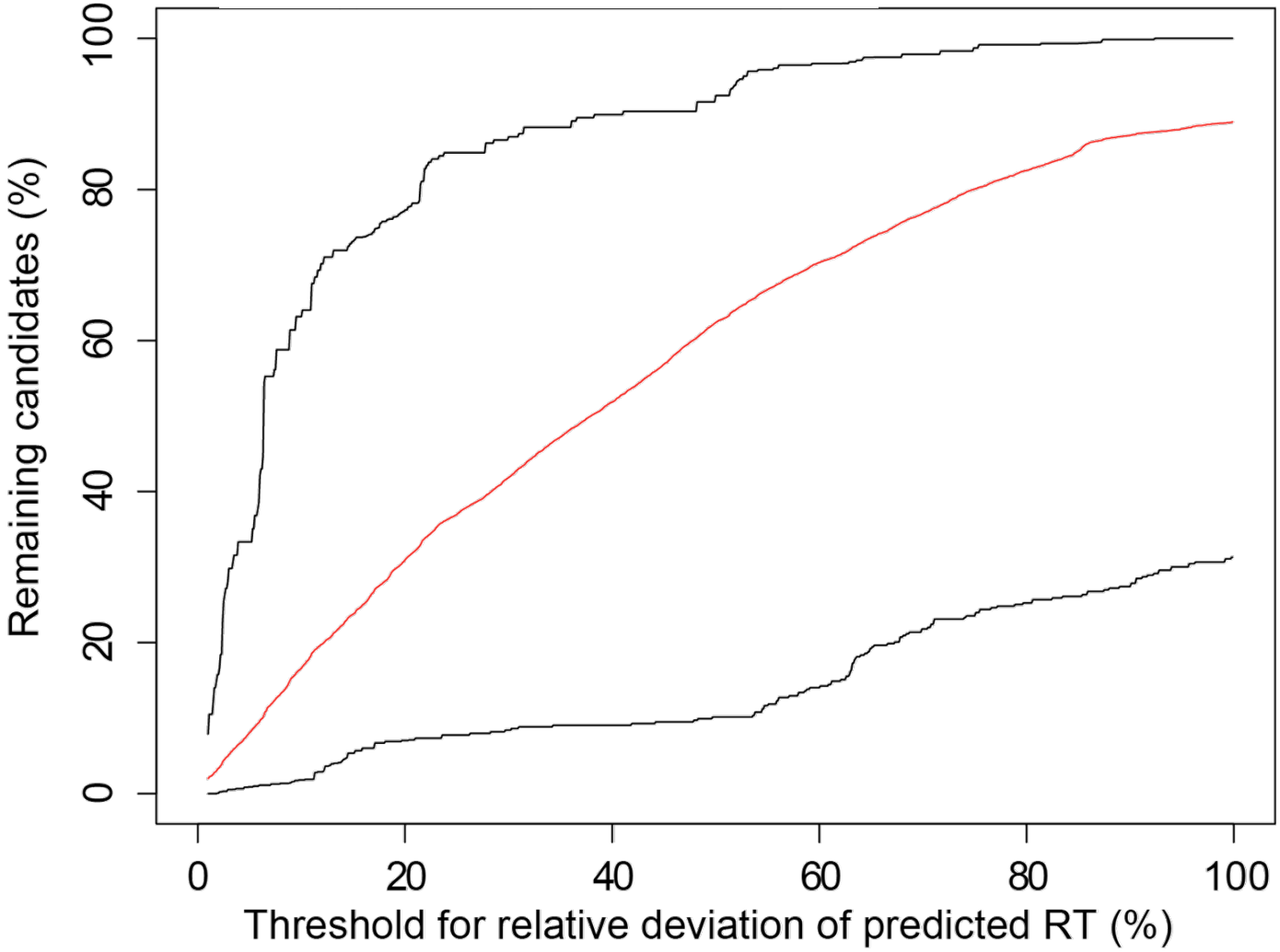
Compound	Name
HMDB00205	Phenylpyruvic acid
HMDB01713	m-Coumaric acid
HMDB02035	4-Hydroxycinnamic acid
HMDB02641	2-Hydroxycinnamic acid

"Piled Higher and Deeper" by Jorge Cham www.phdcomics.com

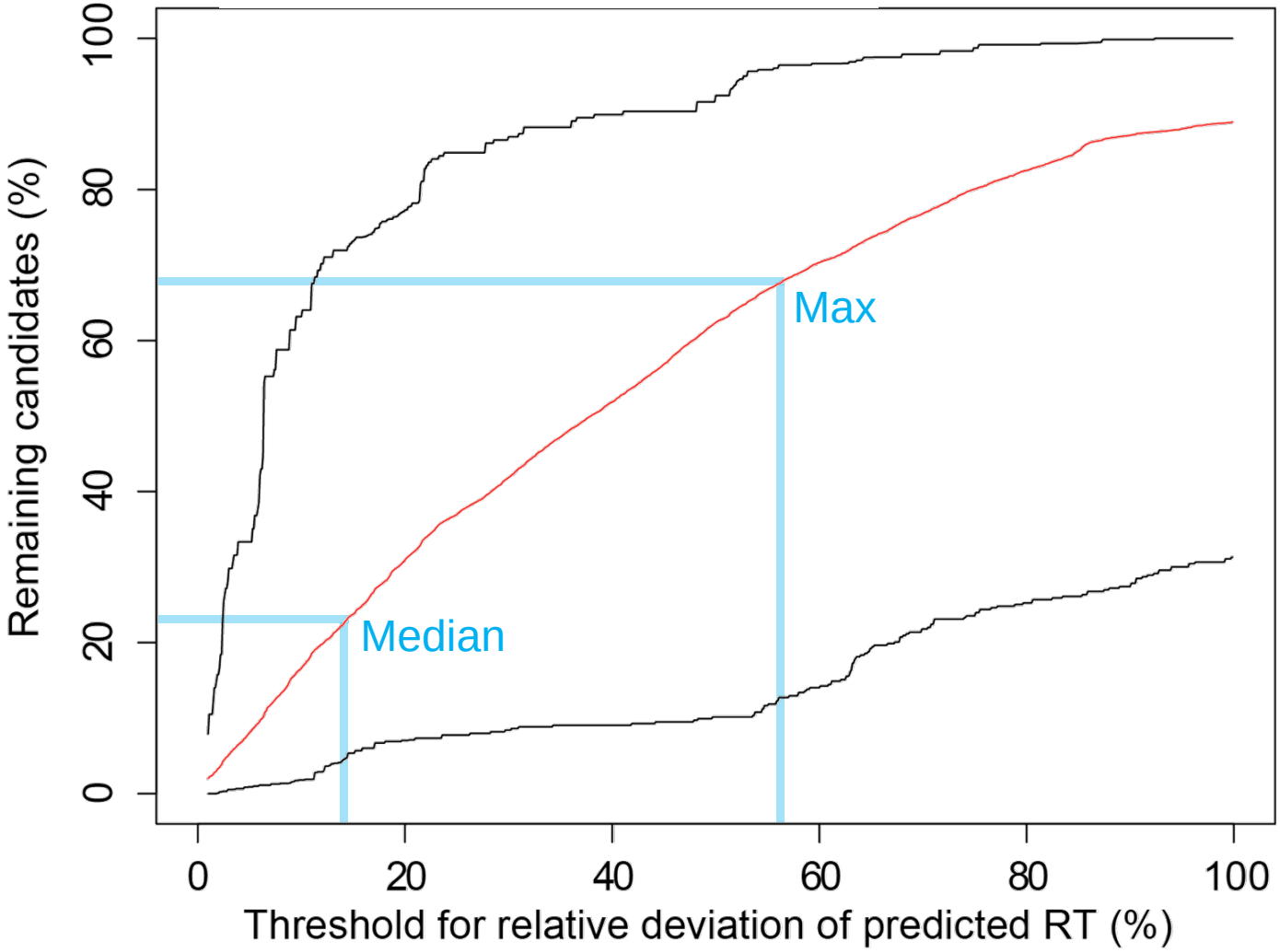
Accuracy of PredRet predictions



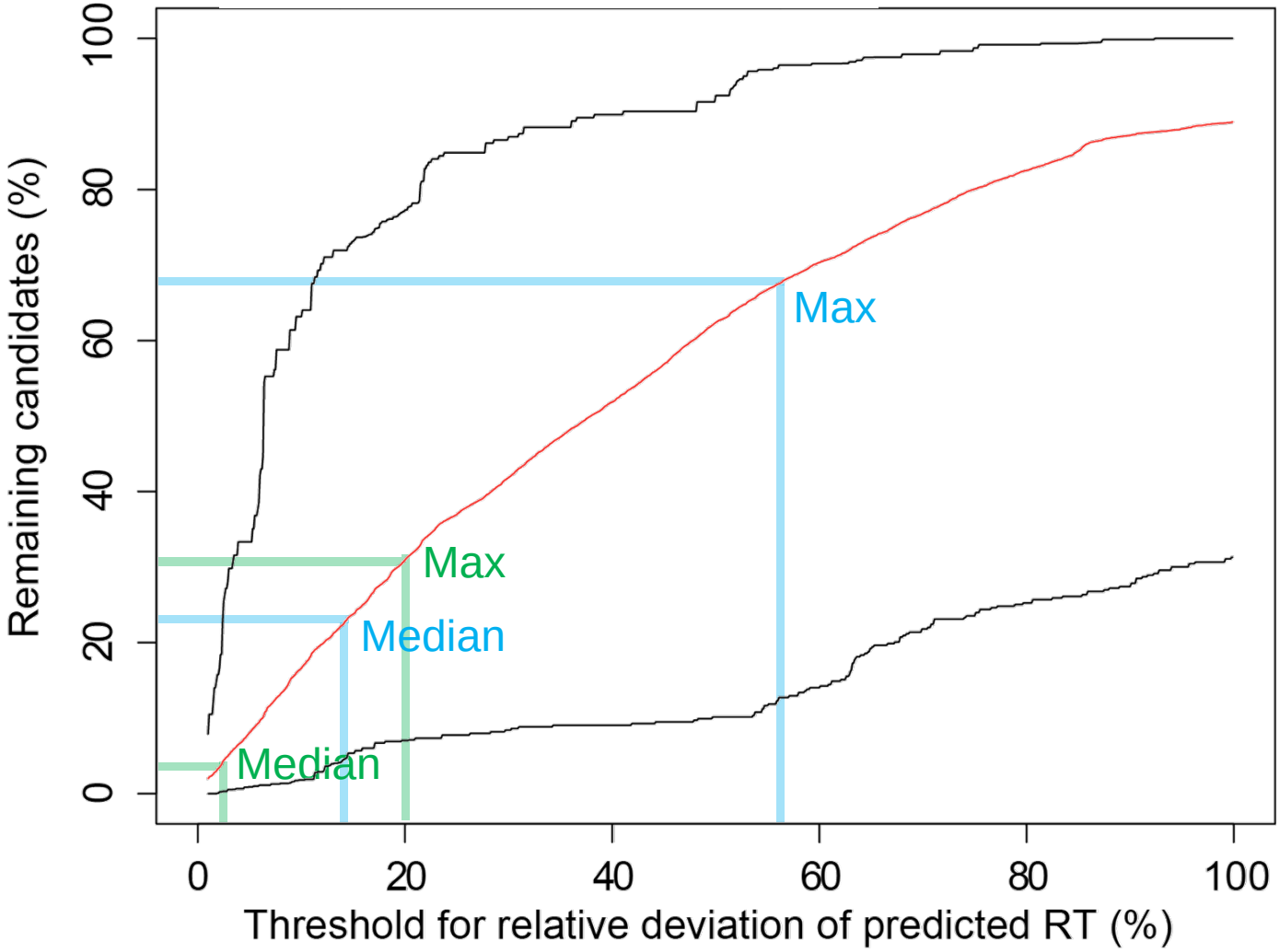
Candidates after filtering as a function of prediction accuracy

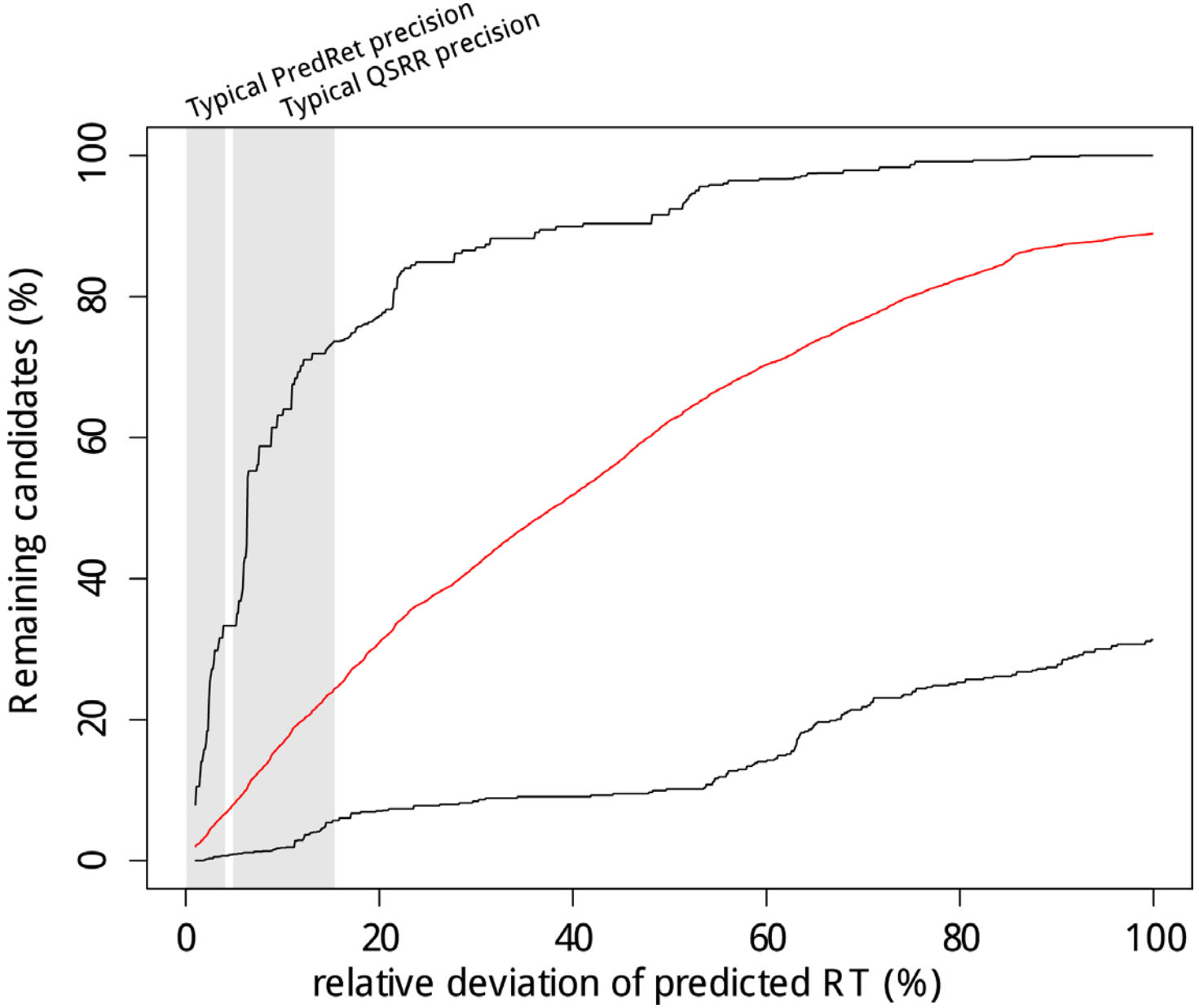


Candidates after filtering as a function of prediction accuracy



Candidates after filtering as a function of prediction accuracy

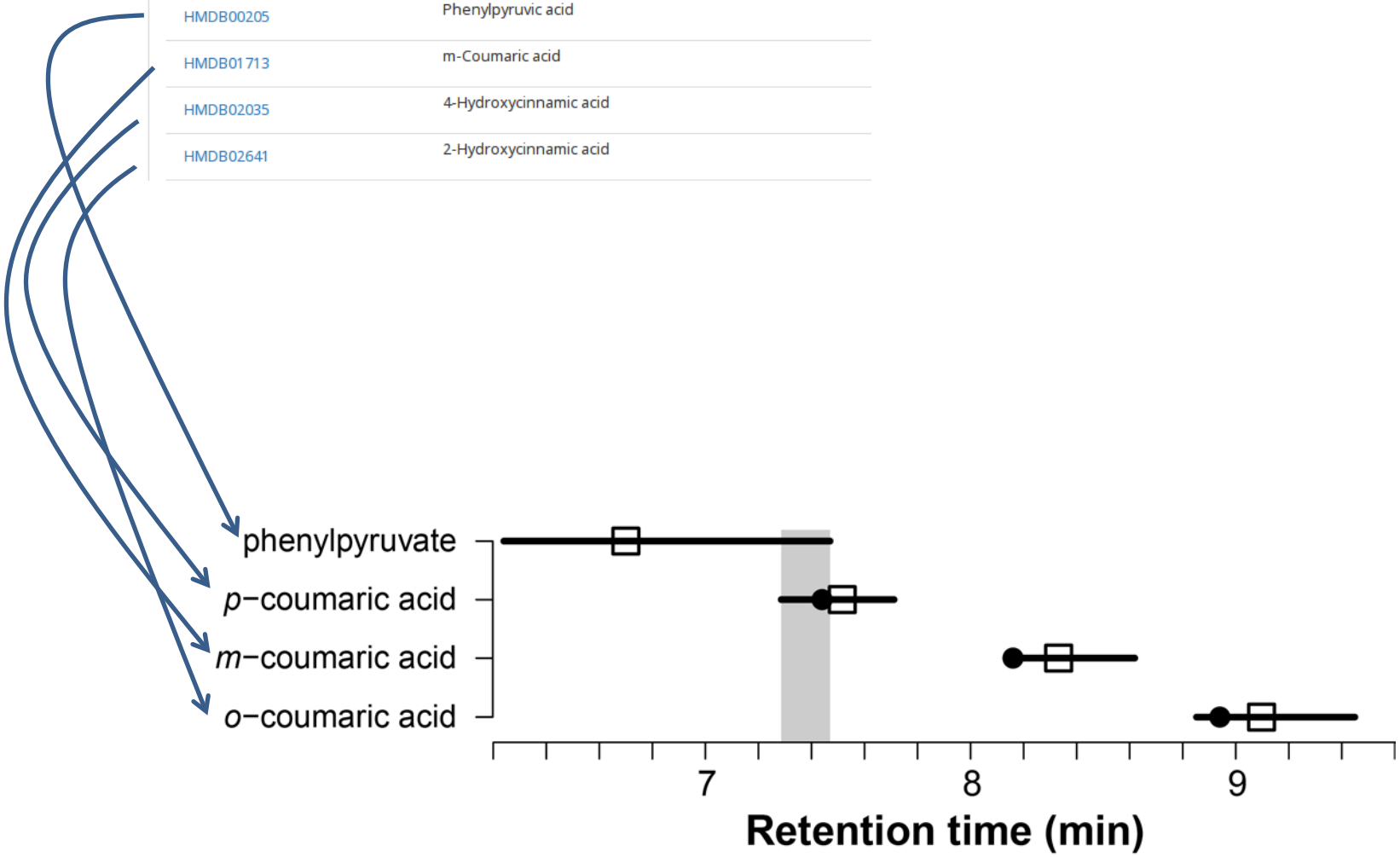



Typical PredRet precision
Typical QSRR precision
Remaining candidates (%)
0
20
40
60
80
100
0
20
40
60
80
100
relative deviation of predicted RT (%)

MS search for 163.04 m/z

Show 10 entries

Compound	Name
HMDB00205	Phenylpyruvic acid
HMDB01713	m-Coumaric acid
HMDB02035	4-Hydroxycinnamic acid
HMDB02641	2-Hydroxycinnamic acid

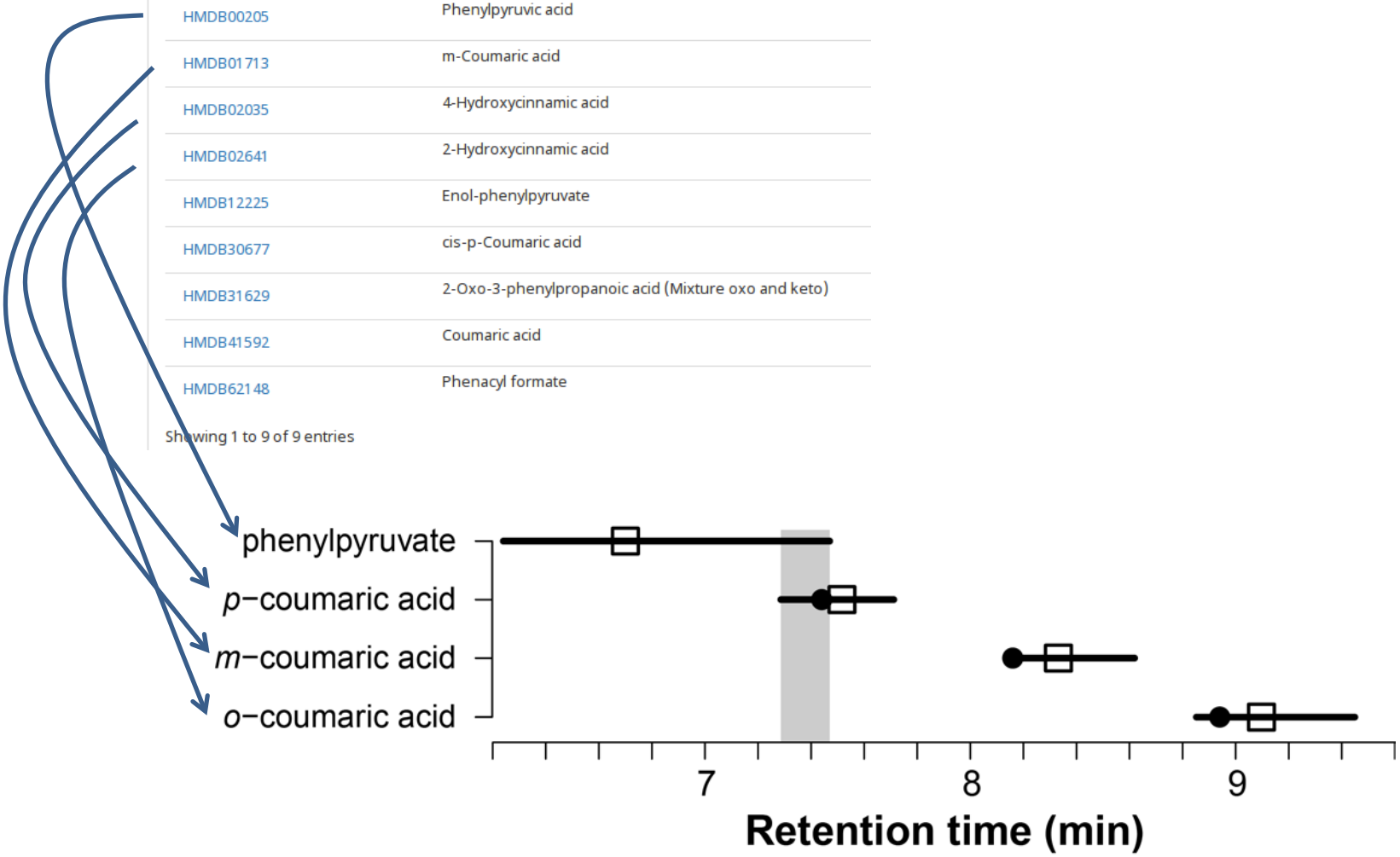


MS search for 163.04 m/z

Show 10 entries

Compound	Name
HMDB00205	Phenylpyruvic acid
HMDB01713	m-Coumaric acid
HMDB02035	4-Hydroxycinnamic acid
HMDB02641	2-Hydroxycinnamic acid
HMDB12225	Enol-phenylpyruvate
HMDB30677	cis-p-Coumaric acid
HMDB31629	2-Oxo-3-phenylpropanoic acid (Mixture oxo and keto)
HMDB41592	Coumaric acid
HMDB62148	Phenacyl formate

Showing 1 to 9 of 9 entries



Conclusions

- PredRet allows sharing of retention time information and across chromatographic systems

Conclusions

- PredRet allows sharing of retention time information and across chromatographic systems
- The accuracy of PredRet can differentiation of isomers

FONDAZIONE EDMUND MACH

THE VELUX FOUNDATIONS
VILLUM FONDEN VELUX FONDEN

Conclusions

- PredRet allows sharing of retention time information and across chromatographic systems
- The accuracy of PredRet can differentiation of isomers

- Compound need to be in the database (*someone* need to record it)

FONDAZIONE EDMUND MACH

THE VELUX FOUNDATIONS
VILLUM FONDEN VELUX FONDEN

Conclusions

- PredRet allows sharing of retention time information and across chromatographic systems
- The accuracy of PredRet can differentiation of isomers

- Compound need to be in the database (*someone* need to record it)
- Chromatography needs to be *similar* (e.g. acidic C18-based)

FONDAZIONE EDMUND MACH

THE VELUX FOUNDATIONS
VILLUM FONDEN VELUX FONDEN

Conclusions

- PredRet allows sharing of retention time information and across chromatographic systems
- The accuracy of PredRet can differentiation of isomers

- Compound need to be in the database (*someone* need to record it)
- Chromatography needs to be *similar* (e.g. acidic C18-based)

- When people use PredRet the database expands

FONDAZIONE EDMUND MACH

THE VELUX FOUNDATIONS
VILLUM FONDEN VELUX FONDEN

Thank you for your attention

Share the RTs!