PCA Sample Identification – Summary Results
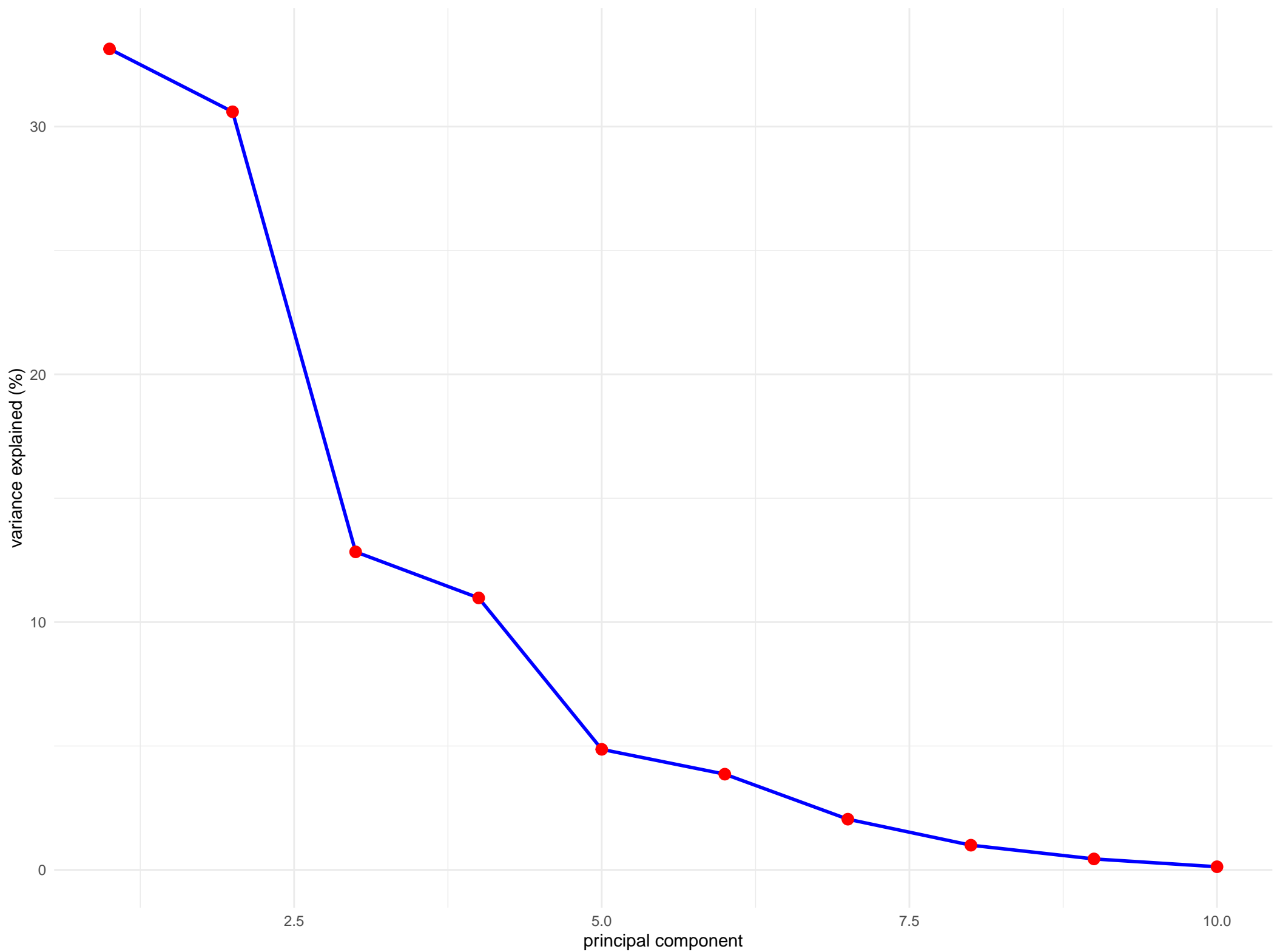
```
ANALYSIS SUMMARY:
o Dataset: 22 samples (5 unknowns + 17 references) | 20 elements analyzed
o Principal Components: PC1: 33.1% | PC2: 30.6% | PC3: 12.8% | Cumulative: 76.6% (>=70% threshold)
o Method: Distance-based identification using 3 principal components
```
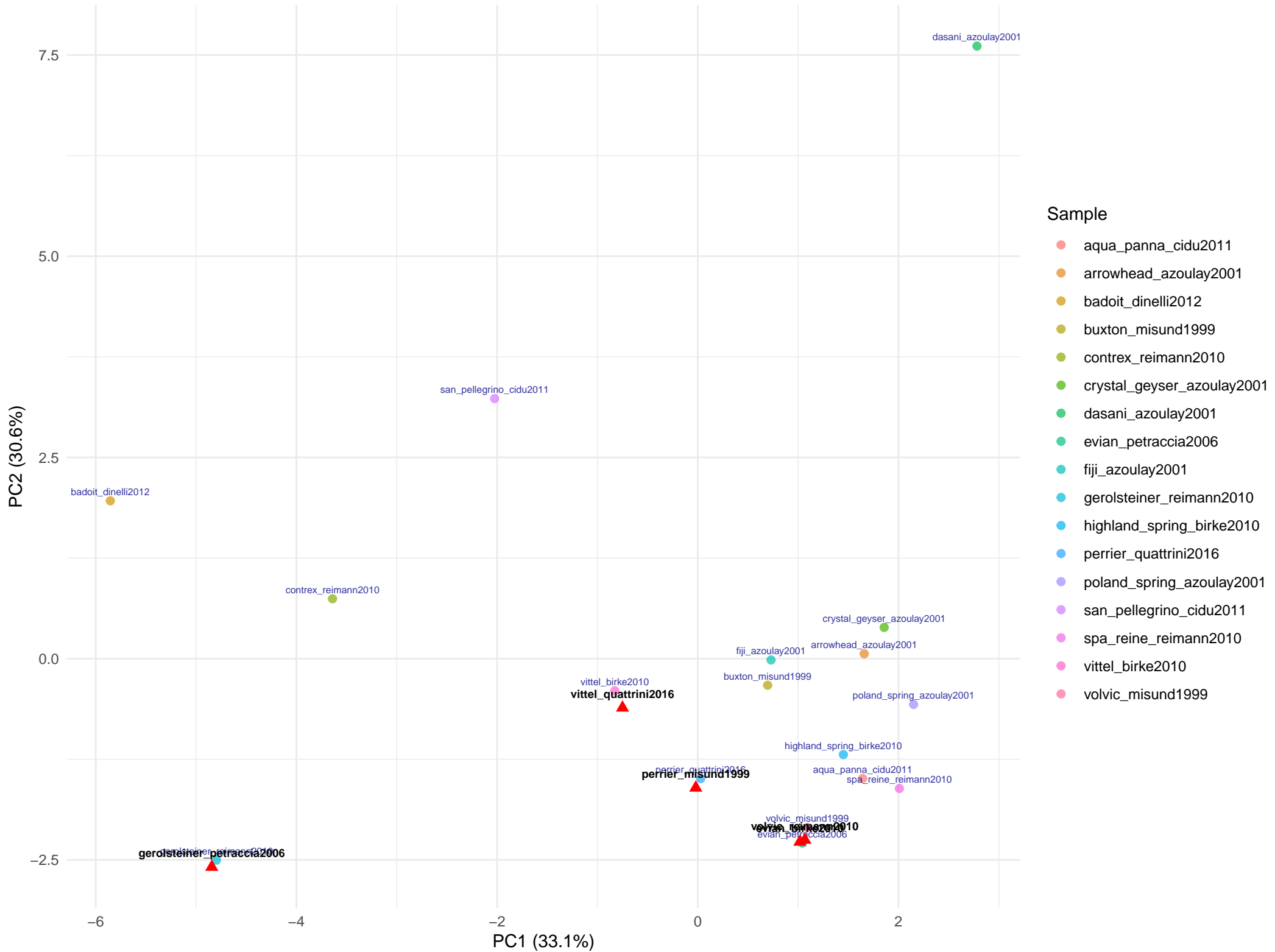
IDENTIFICATION RESULTS

| Sample	Closest Match	Distance	Correlation	Rank
evian_birke2010	evian_petraccia2006	0.042	1.0000	#1
gerolsteiner_petraccia2006	gerolsteiner_reimann2010	0.107	1.0000	#2
volvic_reimann2010	volvic_misund1999	0.180	0.9999	#3
perrier_misund1999	perrier_quattrini2016	0.210	1.0000	#4
vittel_quattrini2016	vittel_birke2010	0.238	1.0000	#5

scree plot – variance explained by principal components

pca scores plot with unknown sample projections

PC2 (30.6%) (y-axis)
PC1 (33.1%) (x-axis)

Labels visible in plot:
- dasani_azoulay2001
- san_pellegrino_cidu2011
- badoit_dinelli2012
- contrex_reimann2010
- crystal_geyser_azoulay2001
- fiji_azoulay2001
- arrowhead_azoulay2001
- buxton_misund1999
- poland_spring_azoulay2001
- vittel_birke2010
- **vittel_quattrini2016**
- highland_spring_birke2010
- perrier_quattrini2016
- **perrier_misund1999**
- aqua_panna_cidu2011
- spa_reine_reimann2010
- volvic_misund1999
- **volvic_reimann2010**
- evian_petraccia2006
- gerolsteiner_reimann2010
- **gerolsteiner_petraccia2006**

Sample (legend):
- aqua_panna_cidu2011
- arrowhead_azoulay2001
- badoit_dinelli2012
- buxton_misund1999
- contrex_reimann2010
- crystal_geyser_azoulay2001
- dasani_azoulay2001
- evian_petraccia2006
- fiji_azoulay2001
- gerolsteiner_reimann2010
- highland_spring_birke2010
- perrier_quattrini2016
- poland_spring_azoulay2001
- san_pellegrino_cidu2011
- spa_reine_reimann2010
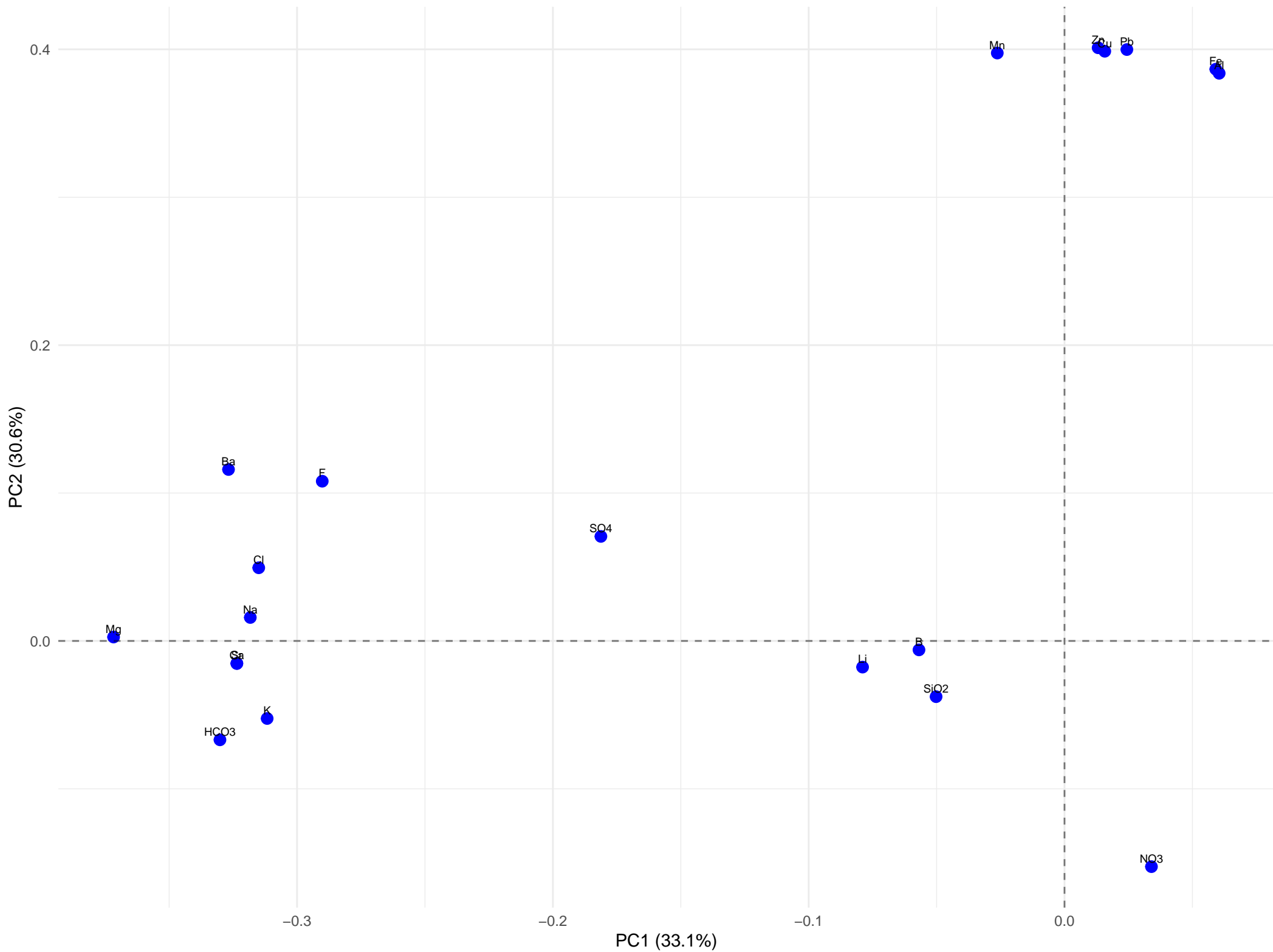- vittel_birke2010
- volvic_misund1999

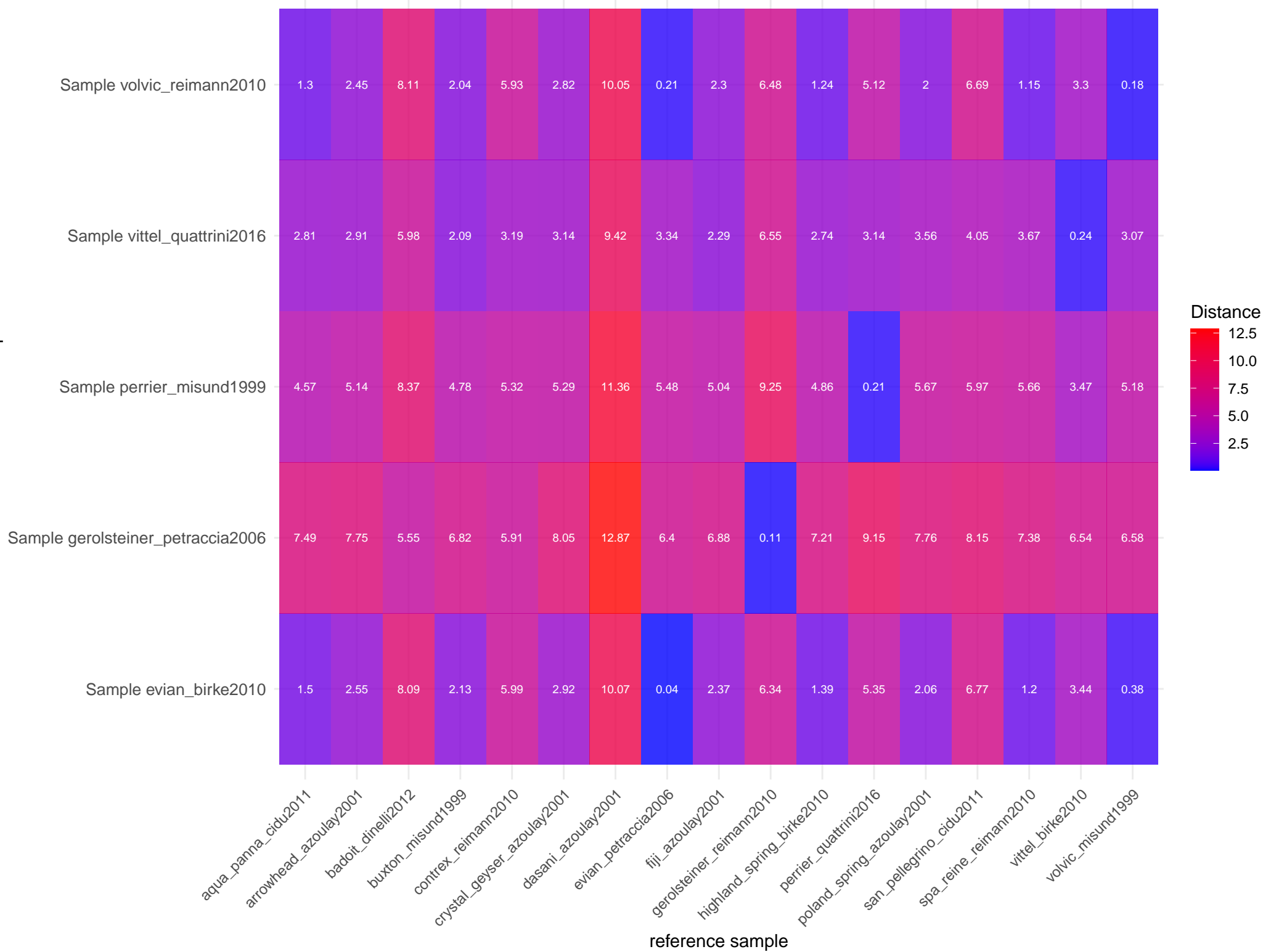pca scores plot: PC1 vs PC3 with unknown sample projections
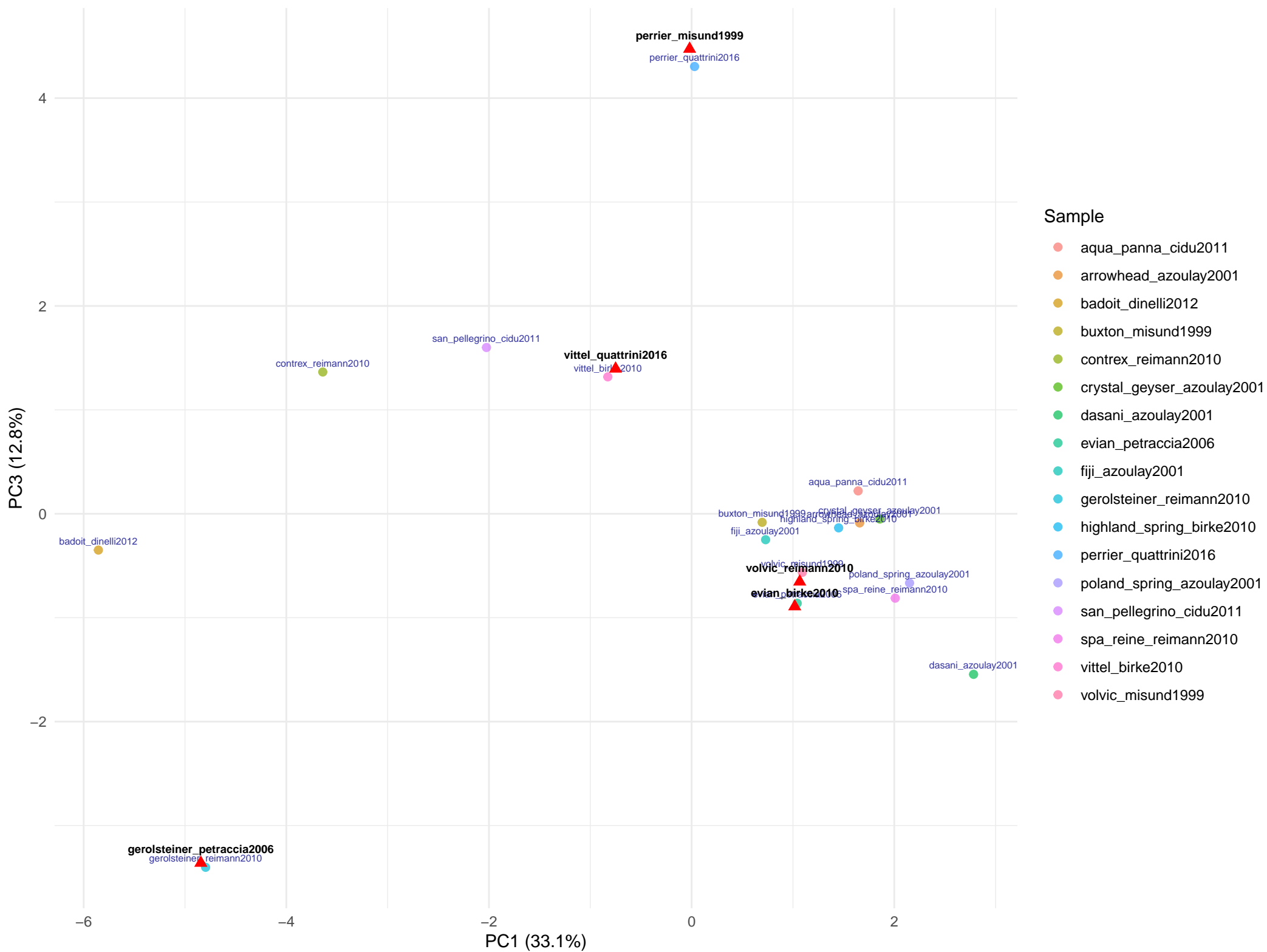
pca scores plot: PC2 vs PC3 with unknown sample projections
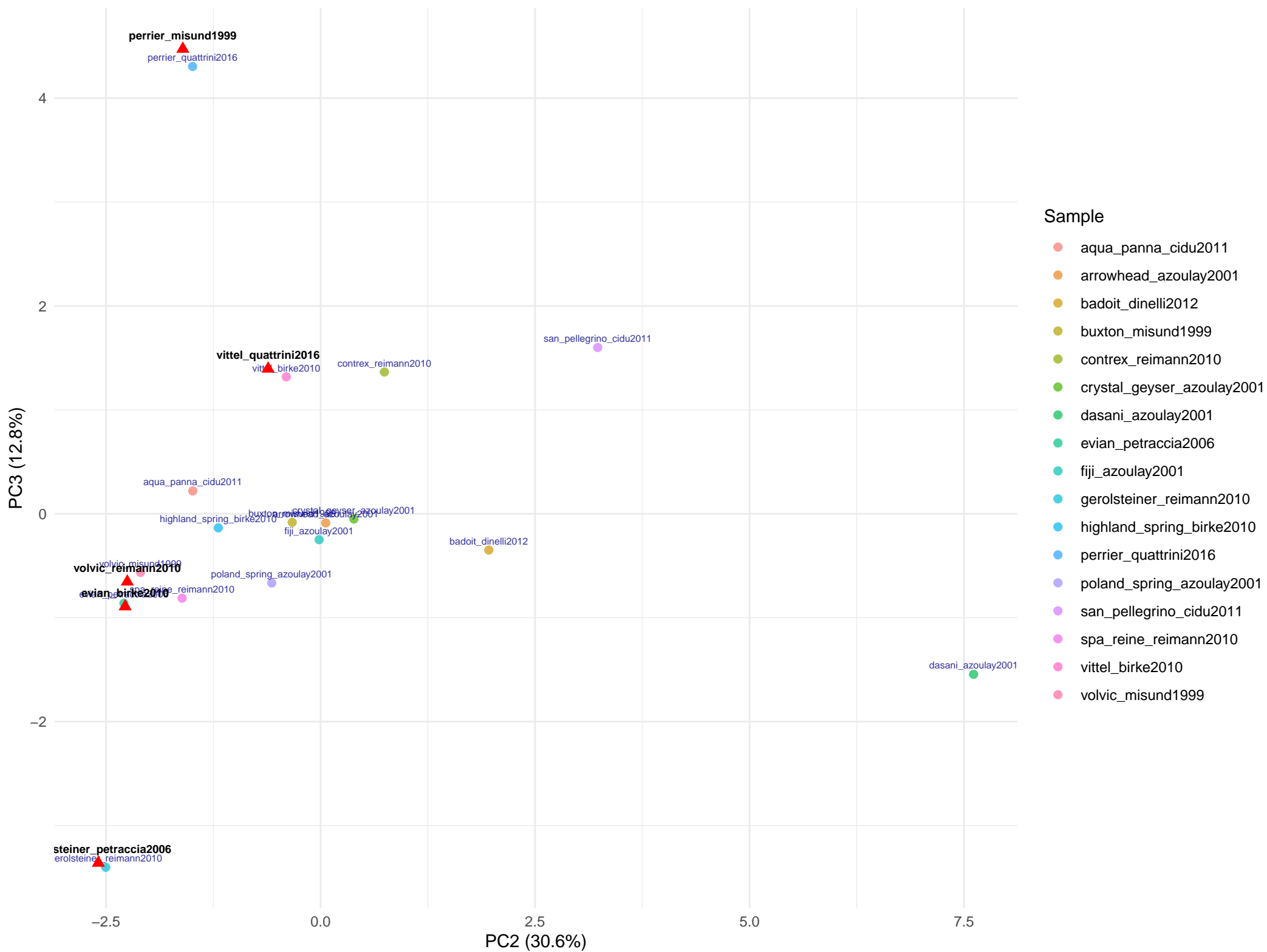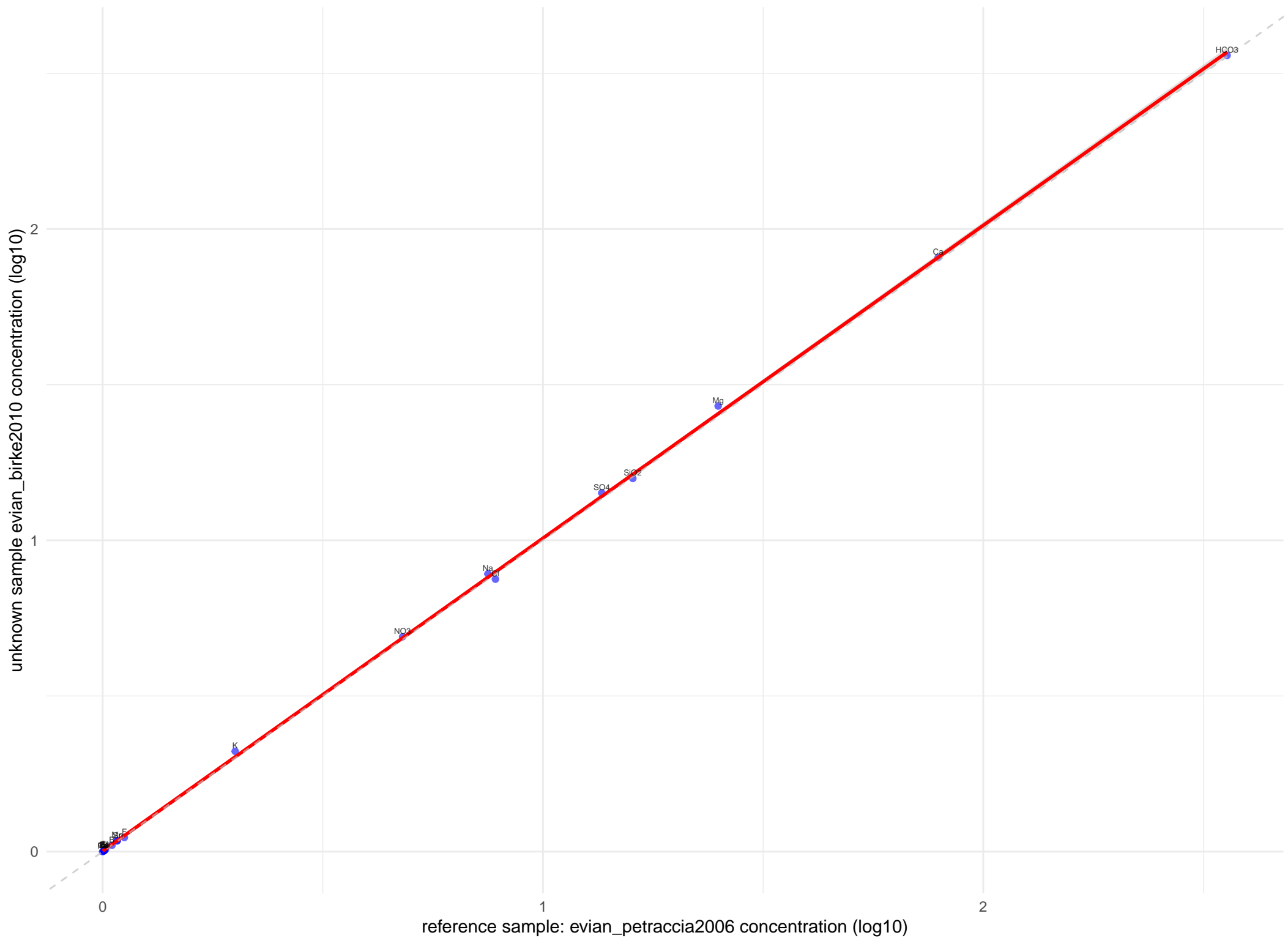
pca loadings plot – top 20 contributing elements

distance matrix – all unknown samples to all reference samples

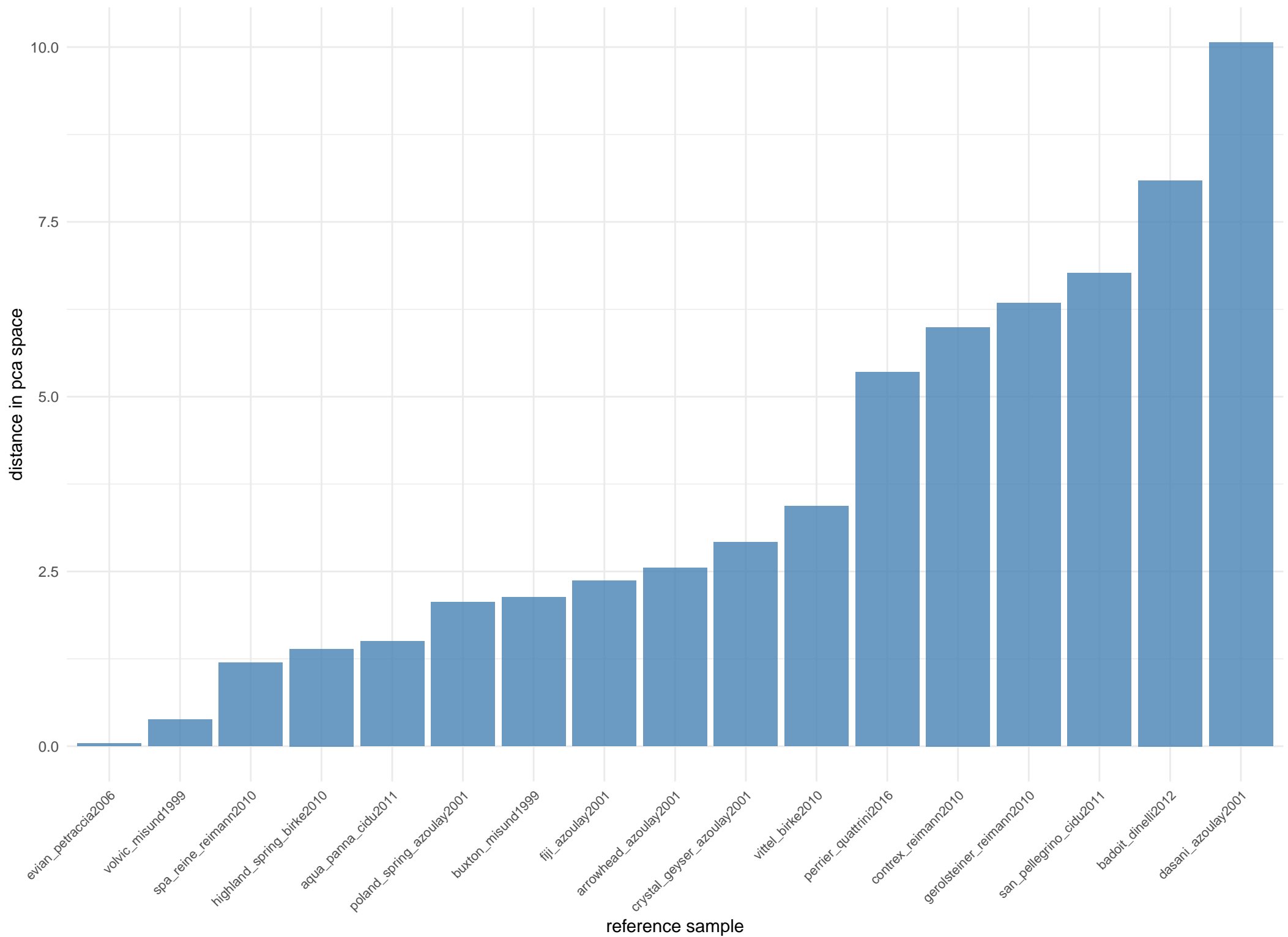**element correlation – sample evian_birke2010 vs evian_petraccia2006**

all 20 elements labeled (log scale) | correlation: r = 1

y-axis: unknown sample evian_birke2010 concentration (log10)

x-axis: reference sample: evian_petraccia2006 concentration (log10)

Labeled points: HCO3, Ca, Mg, SiO2, SO4, Na, NO3, K, F

distance ranking – sample evian_birke2010 to all reference samples

closest match: evian_petraccia2006 | distance: 0.042

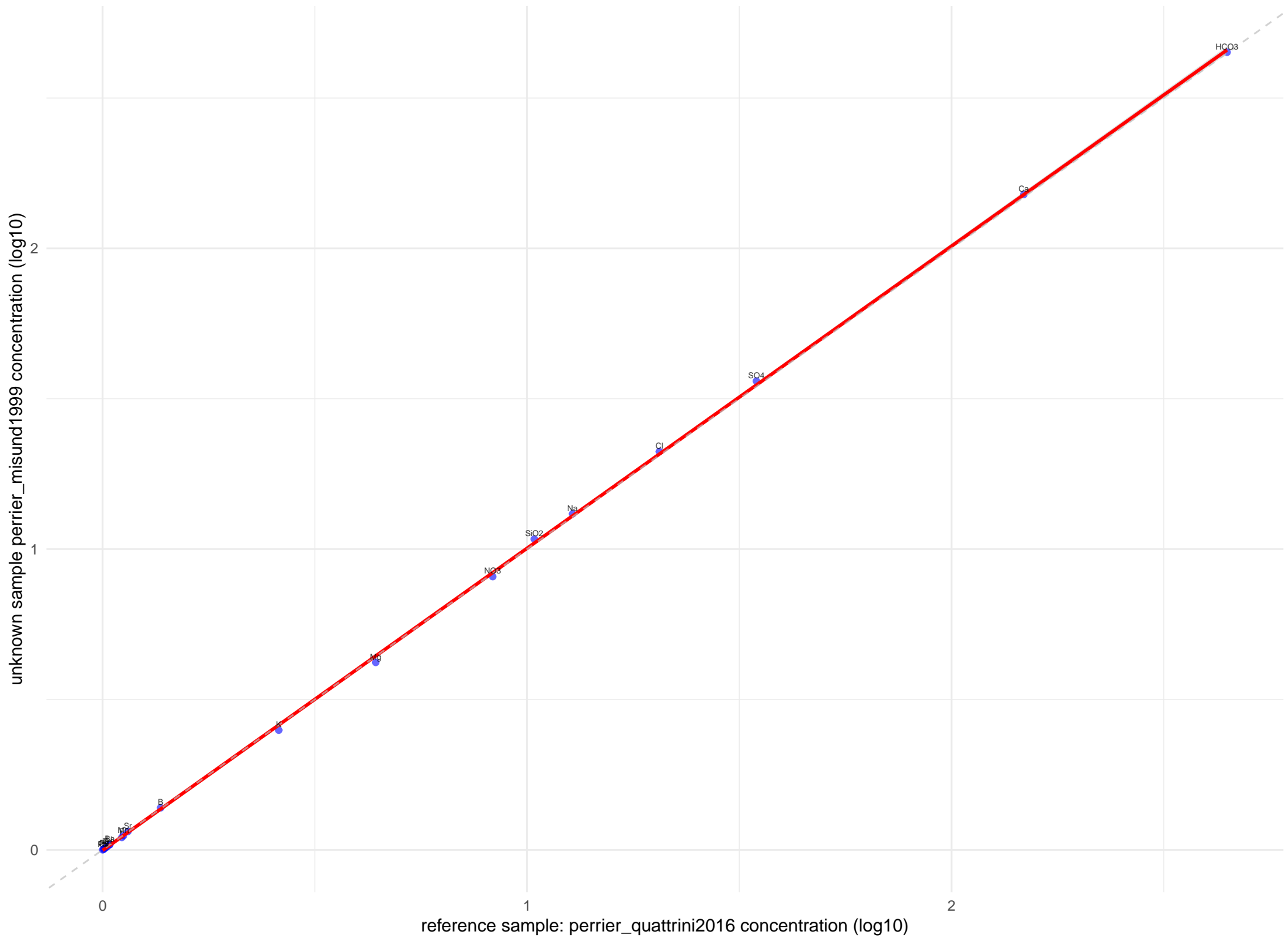**element profile comparison – sample evian_birke2010 vs evian_petraccia2006**

top 20 most variable elements (log scale) | distance: 0.042 | correlation: 1
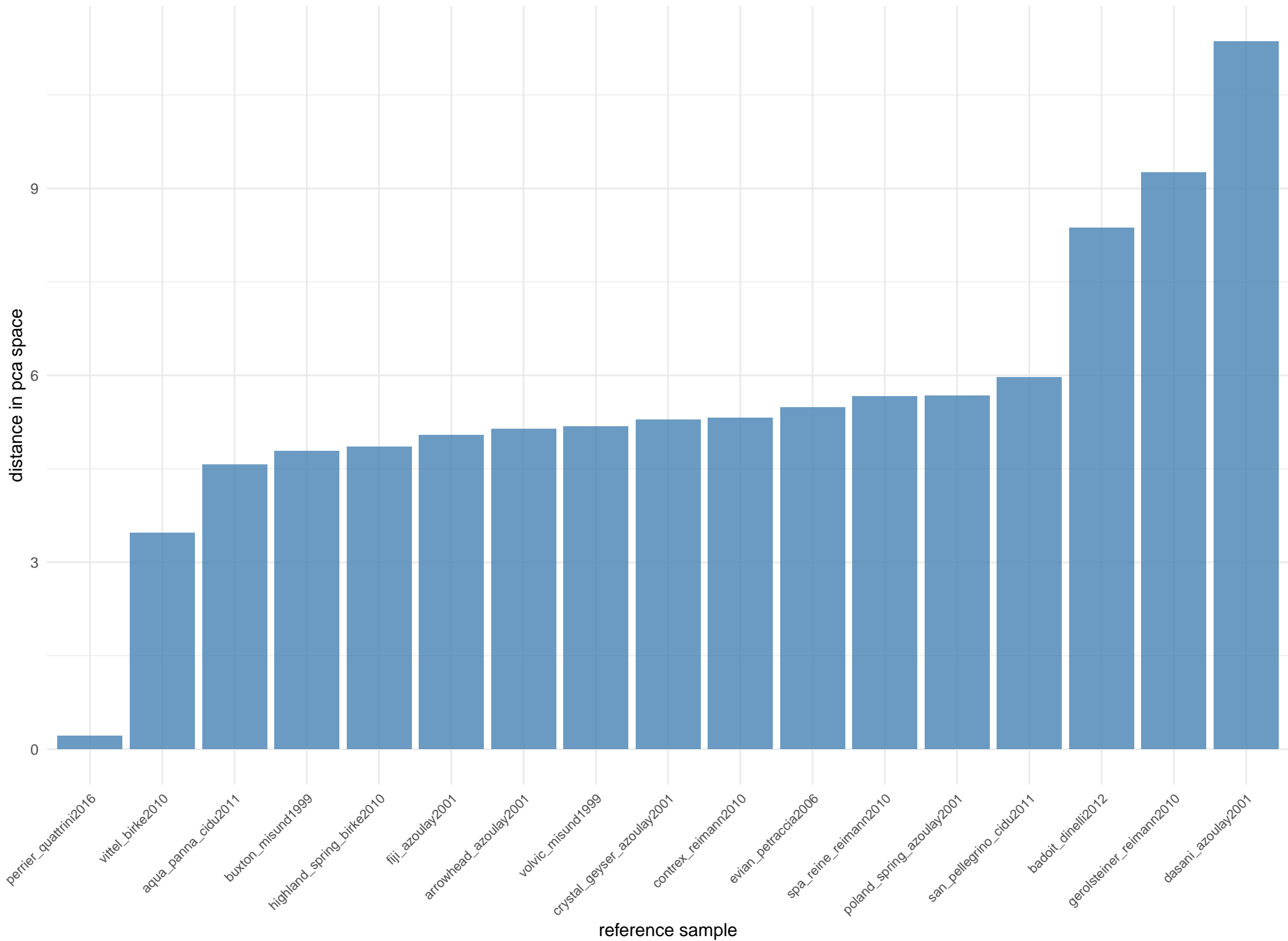
sample ▮ evian_petraccia2006 ▮ Sample evian_birke2010

concentration (log10)

element (top 20 most variable)
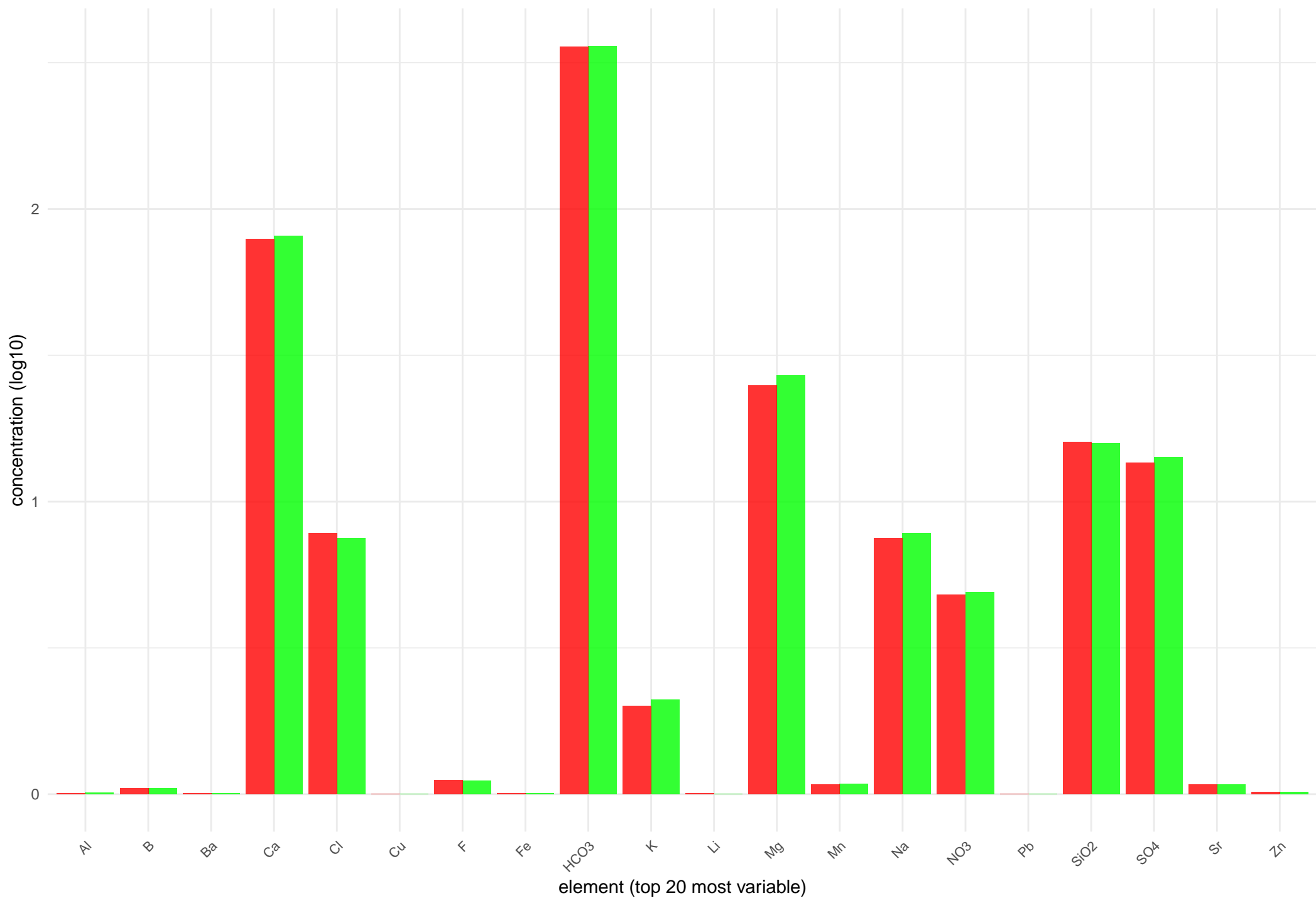
element correlation – sample perrier_misund1999 vs perrier_quattrini2016

all 20 elements labeled (log scale) | correlation: r = 1

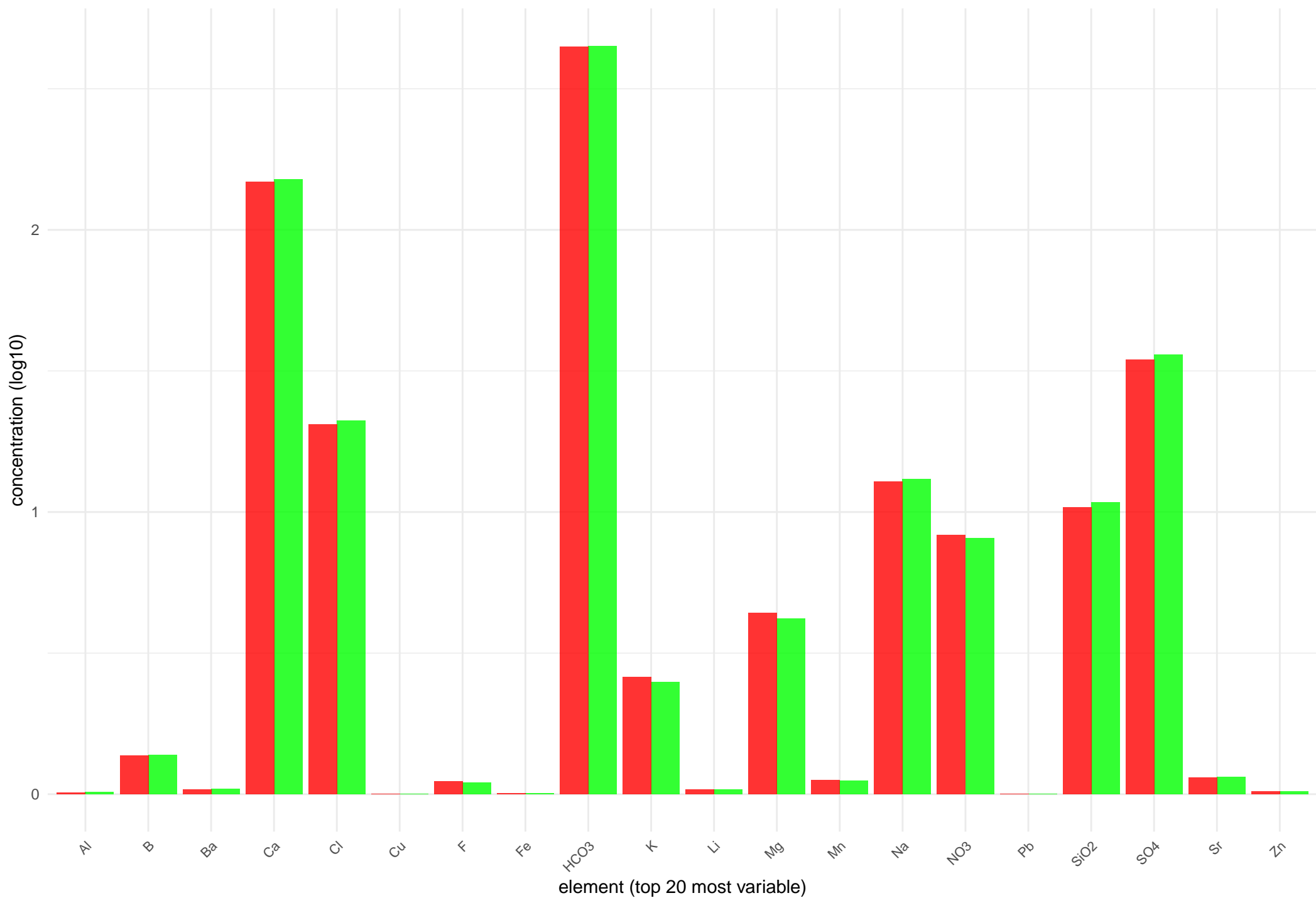**distance ranking – sample perrier_misund1999 to all reference samples**

closest match: perrier_quattrini2016 | distance: 0.21

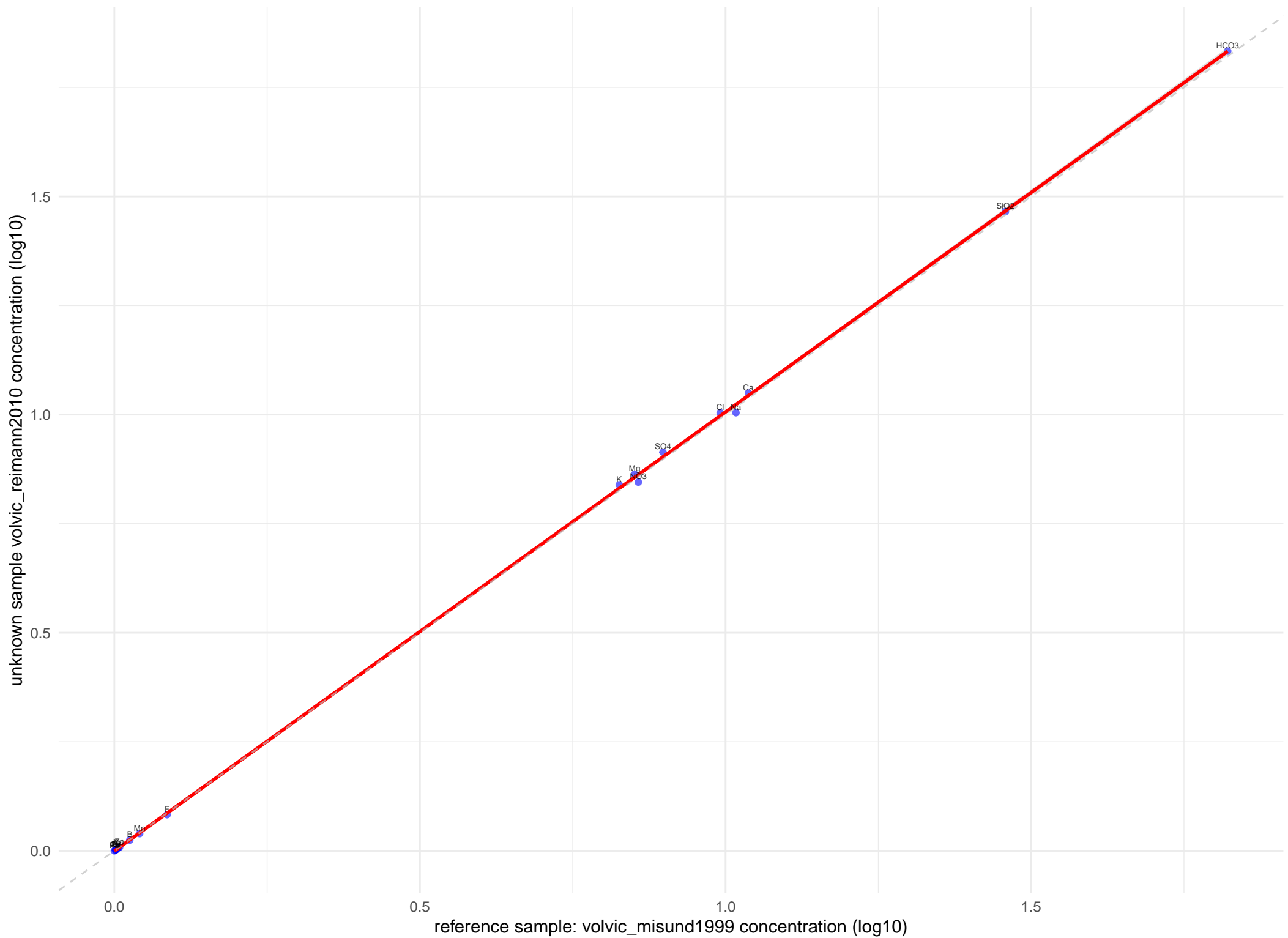**element profile comparison – sample perrier_misund1999 vs perrier_quattrini2016**

top 20 most variable elements (log scale) | distance: 0.21 | correlation: 1

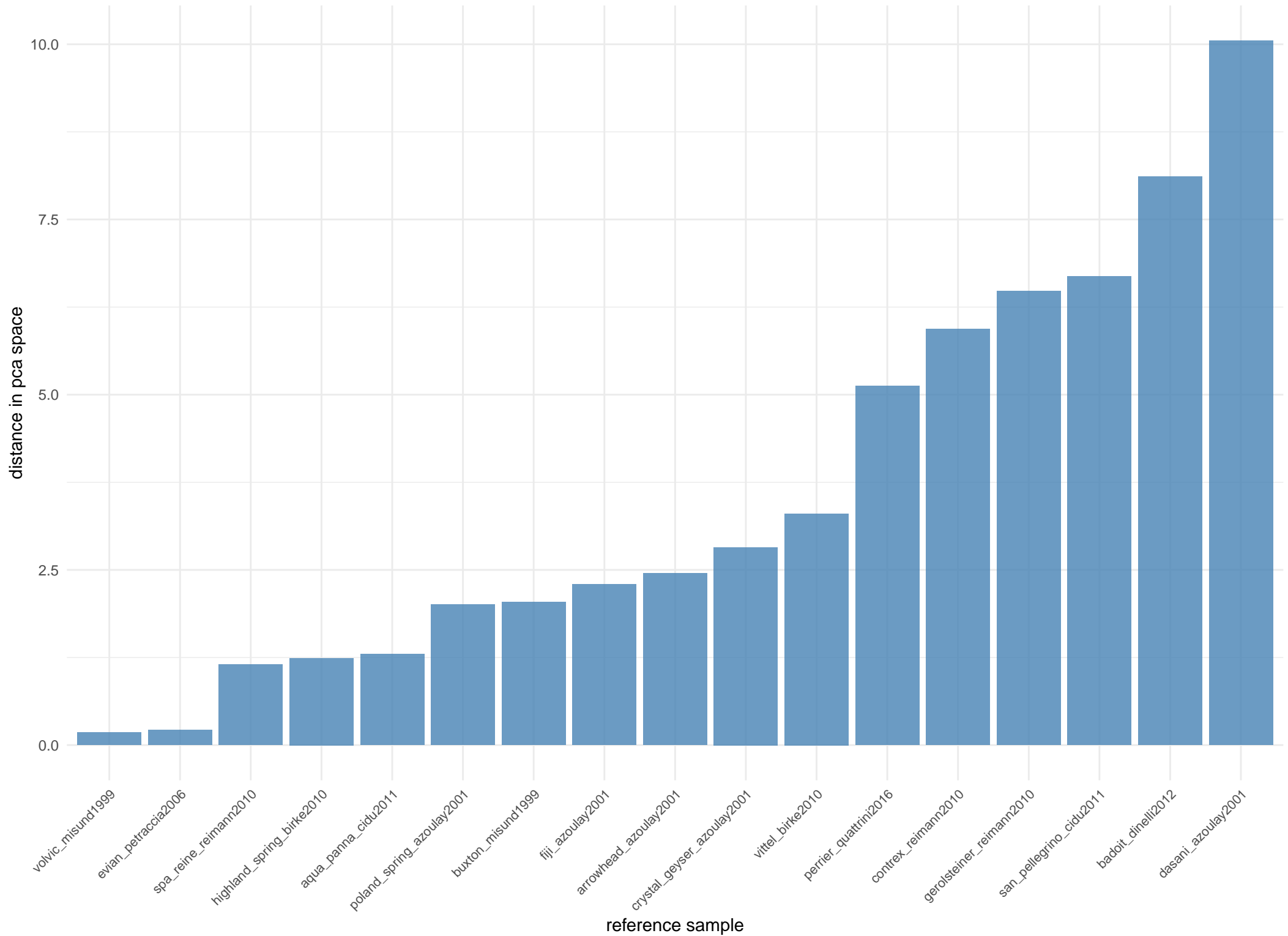sample ▮ perrier_quattrini2016 ▮ Sample perrier_misund1999

element correlation – sample volvic_reimann2010 vs volvic_misund1999

all 20 elements labeled (log scale) | correlation: r = 0.9999

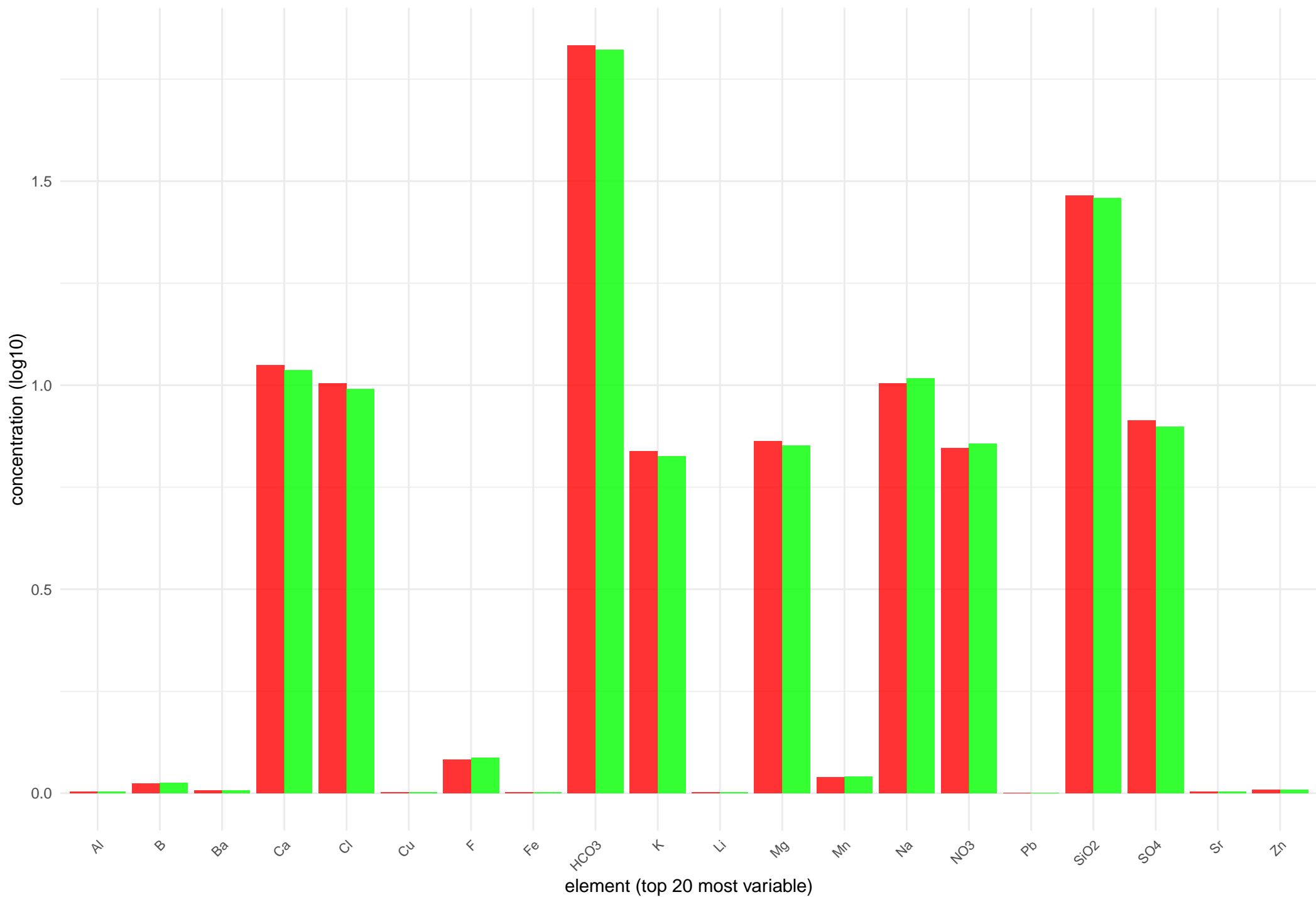**distance ranking – sample volvic_reimann2010 to all reference samples**

closest match: volvic_misund1999 | distance: 0.18

element profile comparison – sample volvic_reimann2010 vs volvic_misund1999

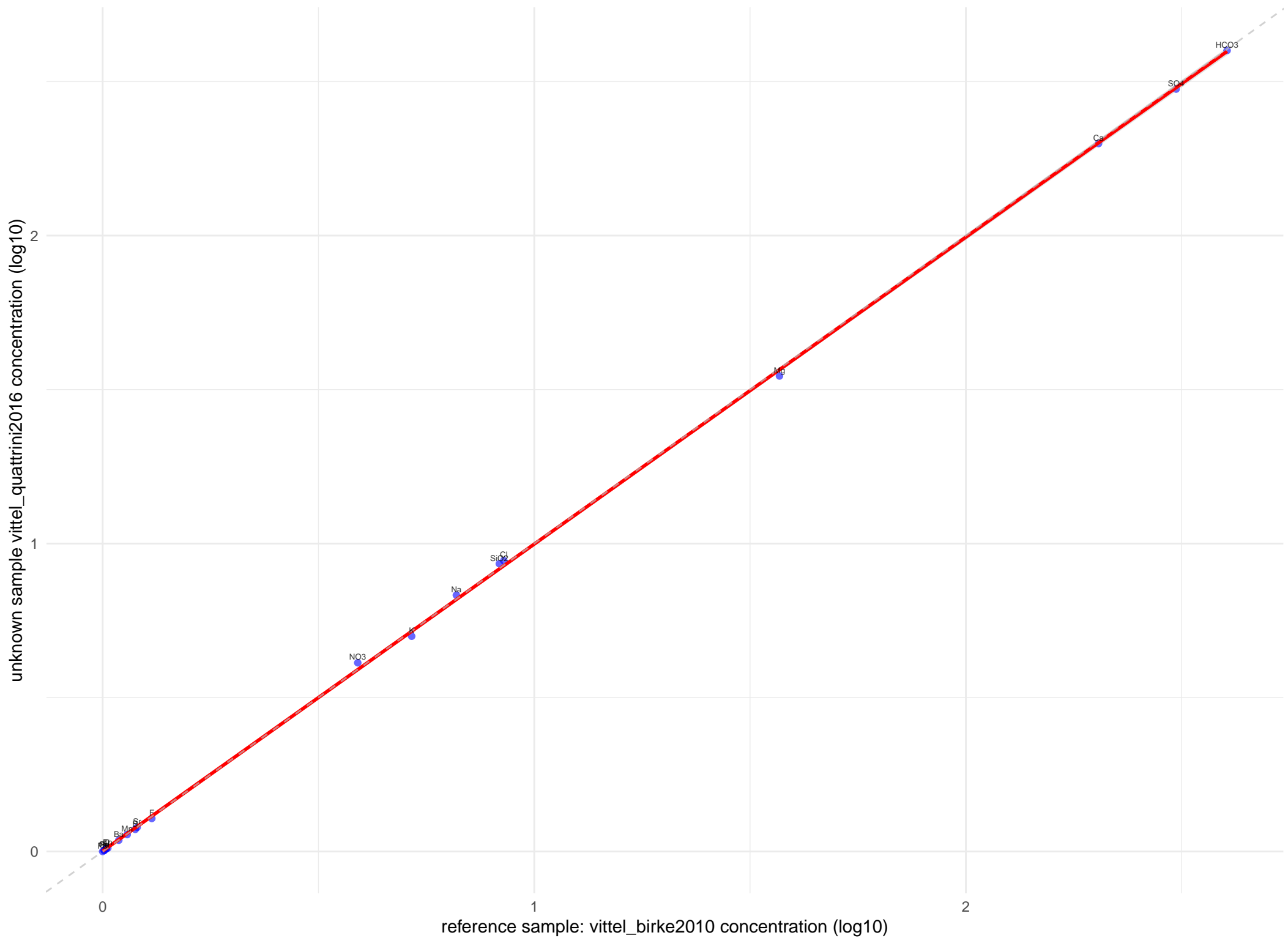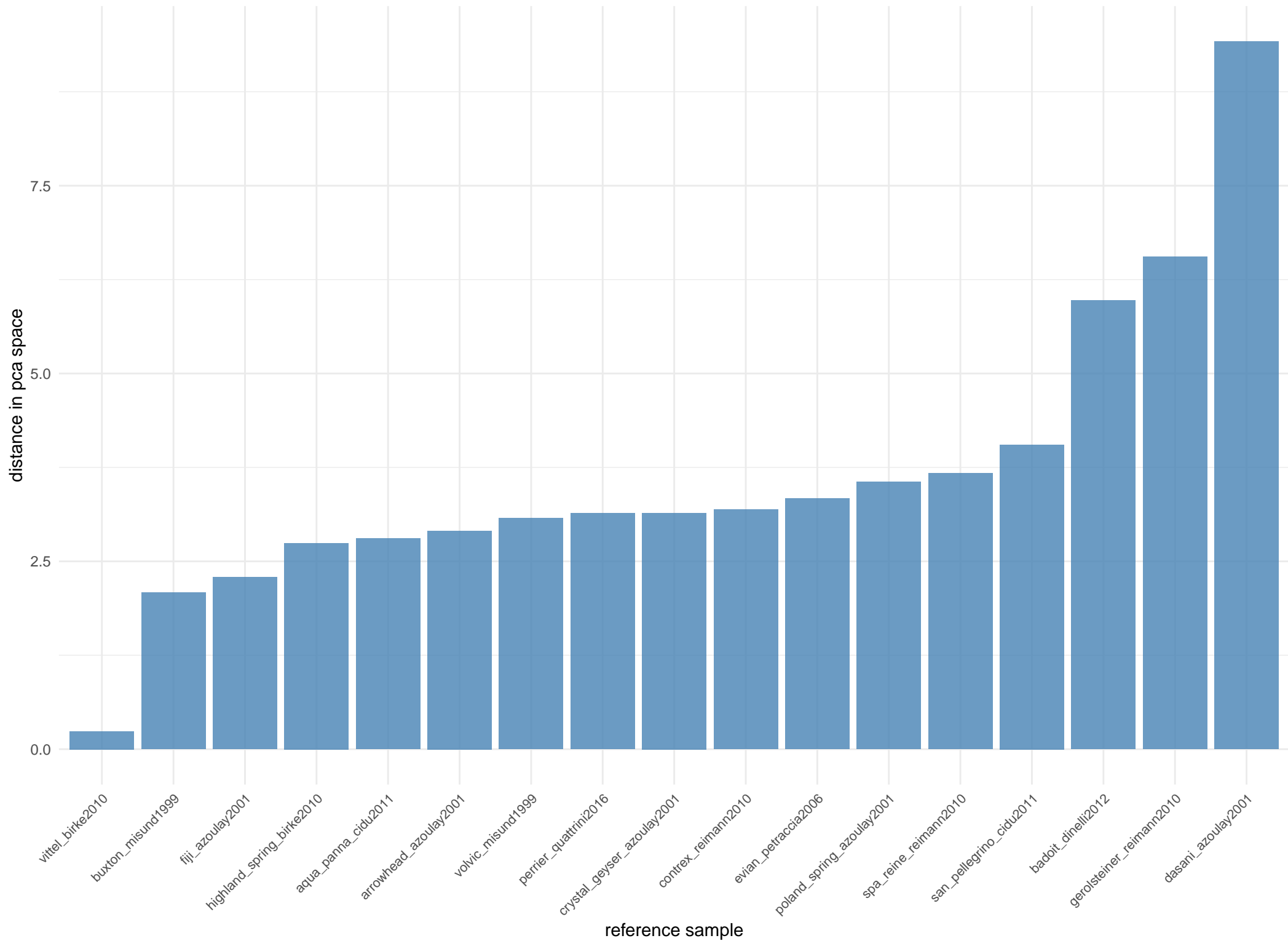top 20 most variable elements (log scale) | distance: 0.18 | correlation: 0.9999

element correlation – sample vittel_quattrini2016 vs vittel_birke2010

all 20 elements labeled (log scale) | correlation: r = 1

x-axis: reference sample: vittel_birke2010 concentration (log10)

y-axis: unknown sample vittel_quattrini2016 concentration (log10)

Labeled elements: HCO3, SO4, Ca, Mg, Cl, SiO2, Na, K, NO3, F, Sr, Mn, Ba

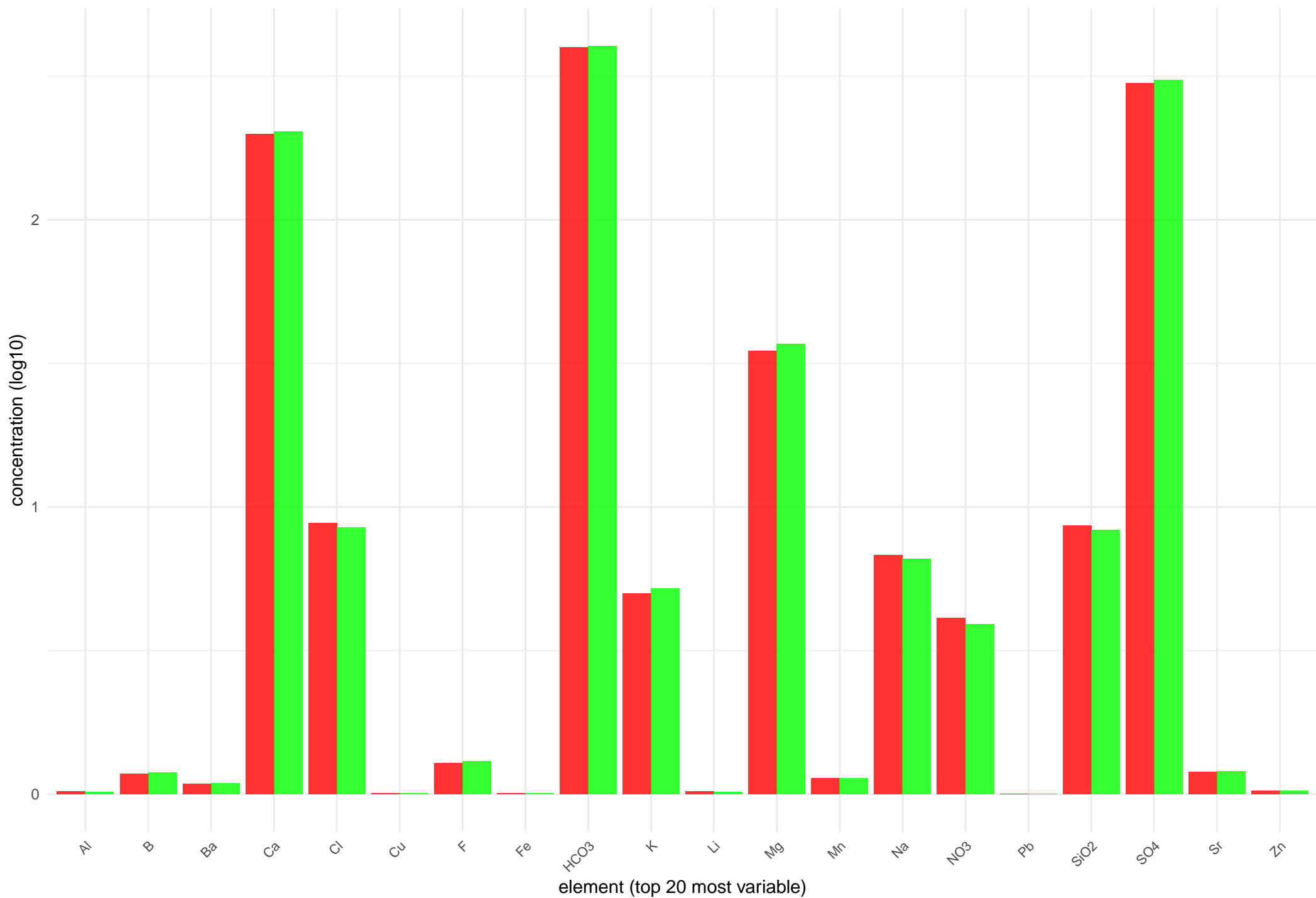**distance ranking – sample vittel_quattrini2016 to all reference samples**

closest match: vittel_birke2010 | distance: 0.238

element profile comparison – sample vittel_quattrini2016 vs vittel_birke2010
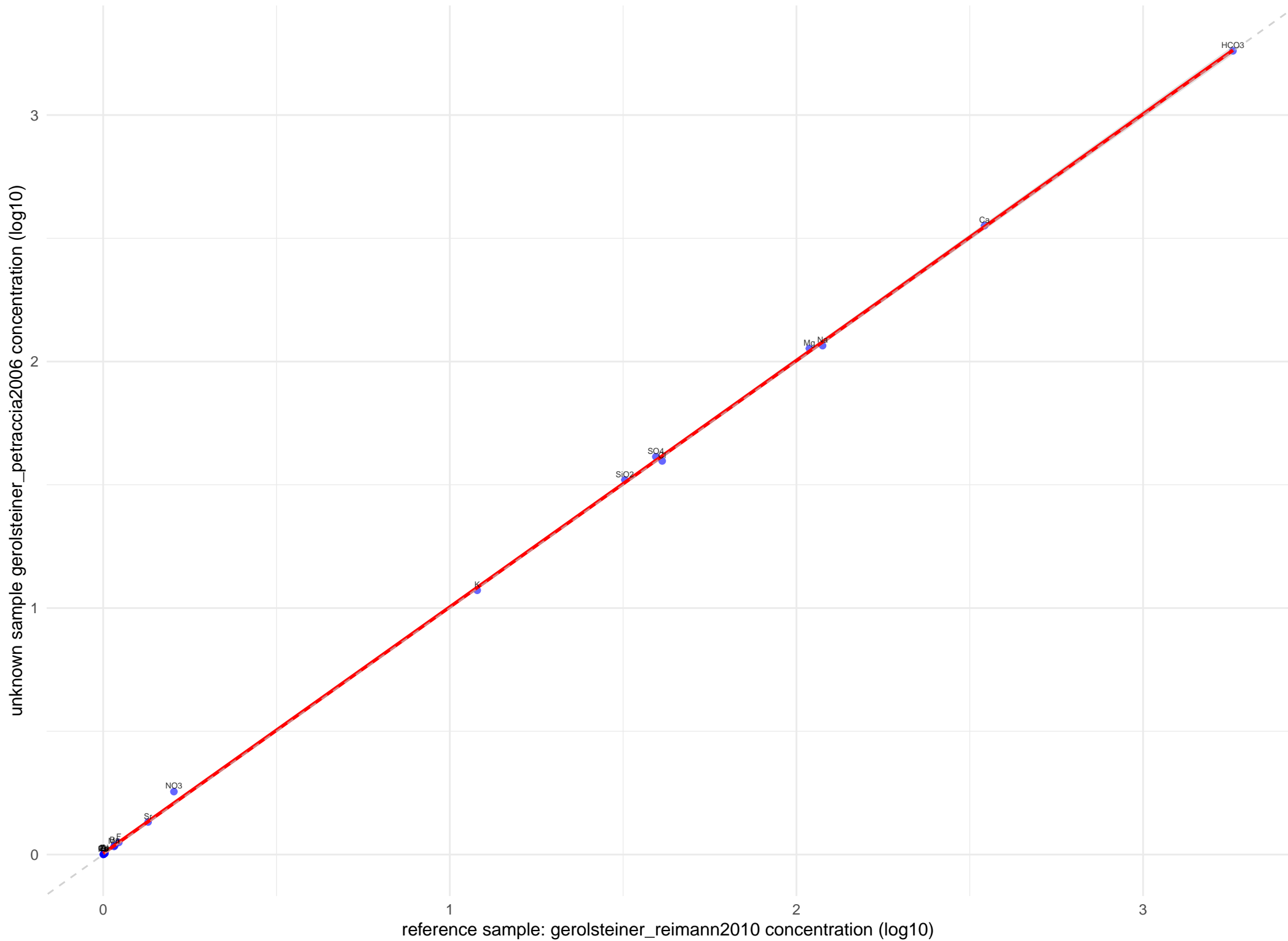
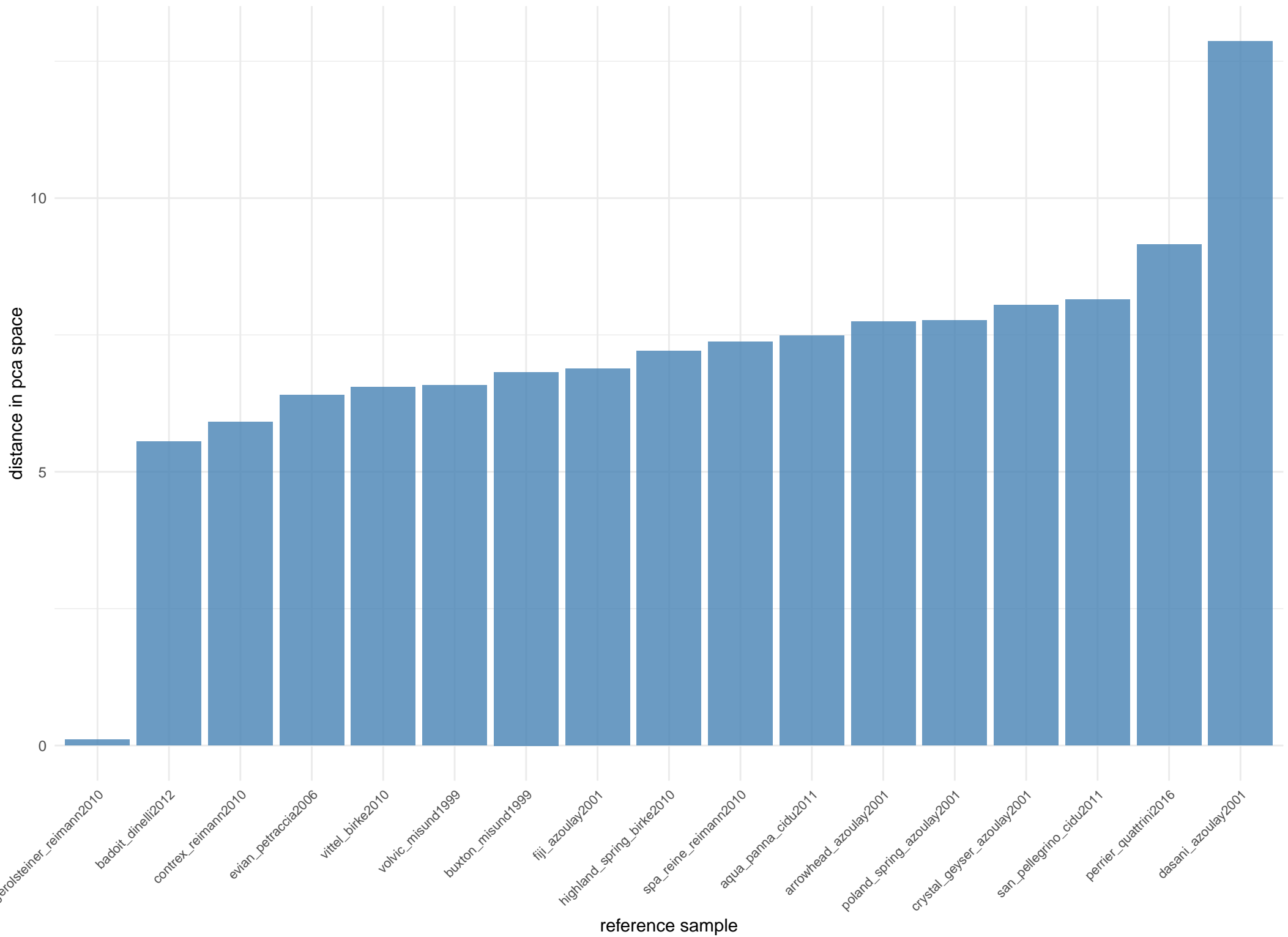top 20 most variable elements (log scale) | distance: 0.238 | correlation: 1

element correlation – sample gerolsteiner_petraccia2006 vs gerolsteiner_reimann2010

all 20 elements labeled (log scale) | correlation: r = 1

Y-axis: unknown sample gerolsteiner_petraccia2006 concentration (log10)

X-axis: reference sample: gerolsteiner_reimann2010 concentration (log10)

distance ranking – sample gerolsteiner_petraccia2006 to all reference samples

closest match: gerolsteiner_reimann2010 | distance: 0.107

element profile comparison – sample gerolsteiner_petraccia2006 vs gerolsteiner_reimann2010

top 20 most variable elements (log scale) | distance: 0.107 | correlation: 1

sample ■ gerolsteiner_reimann2010 ■ Sample gerolsteiner_petraccia2006