



Simple MassChroQ User Manual

Free and Open Source Mass Chromatogram
Quantification Software

version: 2.4.30

Filippo Rusconi
Edlira Nano
Michel Zivy

Benoît Valot
Olivier Langella

Contents

Introduction	3
Input format	3
Output format	5
MSrun retention time alignment	6
Peak quality code	6
Match between run difference with legacy MassChroQ	7

Introduction

SimpleMassChroQ is designed to extract ion current of a list of peptide and measure the area under the curve on a single run. No match between run.

Input format

input file format

```
{
  "project_parameters": {
    "i2MassChroQ_VERSION": {
      "category": 1,
      "value": "1.0.18"
    },
    "cvparam_SpectraDataFileFormat": {
      "category": 5,
      "value": "MS:1000544"
    },
    "cvparam_SpectraDataSpectrumIDFormat": {
      "category": 5,
      "value": "MS:1000824"
    },
    "AnalysisSoftware_name": {
      "category": 2,
      "value": "X!Tandem"
    },
  },
  "masschroq_methods": {
    "alignment_method": {
      "ms2_tendency": 10,
      "ms1_smoothing": 3,
      "ms2_smoothing": 5,
    },
    "quantification_method": {
      "extraction": {
        "integration": "max",
        "precision": {
          "unit": "ppm",
          "up": 10,
          "down": 10,
        }
      },
    },
    "prefilter": {},
    "detection": {
      "type": "zivy",
      "meanfilter": 1,
      "minmax": 3,
      "maxmin": 2,
      "threshold_on_max": 5000,
      "threshold_on_min": 3000,
    }
  }
}
```

```
},

"identification_data": {
  "msrun_list": {
    "msrunb19": {
      "file": "/gorgone/pappso/data_extraction_pappso/
mzXML/20120906_balliau_extract_1_B08_teal-5.mzXML"
    }
  },
  "protein_list": {
    "protala1": {
      "description": "P02769|ALBU_BOVIN SERUM ALBUMIN PRECURSOR",
      "sequence": "SLTNDWEDHLAVK",
    },
  },
},

"peptide_list": {
  "pepala1": {
    "proformat": "SLTNDWEDHLAVK",
    "proteins": ["protala1"],
    "label_list": { "light": {"proformat": "SLTNDWEDHLAVK"}
  },
},
},

"msrunpeptide_list": {"msruna1": {
  "peptide_obs": { "pepala1":
    [{
      "scan_index": 2345,
      "label": "light",
      "precursor": {
        "charge": 2,
        "mz": 2456.45,
        "intensity": 4580,
        "rt": 345.67,
      }
    },
  ],
},
},
},

"action": {
  "match_between_run": true,
  "isotope_minimum_ratio": 0.9,
  "group_list": { "g1": ["msruna1", "msruna2", "msruna3"]
  },
  "quantify_group": {"g1": {
    "alignment_reference": "msruna1"
  }}
}
```

```
}
}
```

CSV input file format

peptide proformat, charge, retention time

SLTNDWEDHLAVK, 2, 853.78

SLTNDWEDHLAVK, 3, 853.78

Output format

Wouldn't be nice to try CBOR ? CBOR is supported in QT6

It could be something liket that :

output file format

```
{ "quantification_data" : [
  {
    "quantify_id": "q1",
    "group_id": "g1",
    "first_pass": { "msruna1": "qr_data_block", "msruna2": "qr_data_block" },
    "second_pass": { "msruna1": "qr_data_block", "msruna2": "qr_data_block" }
  }
]
}
```

QrDataBlock element

QrDataBlock stands for **quantification run data block** it handles quantification data for a single MSrun.

qr_data_block element format

```
{ "msruna1" : {
  "msrun" : { "id": "msruna1", "filename": "/gorgone/truc", "sample":
"échantillon" },
  "retention_time_correction": { "original": [1,2,3,4.5], "aligned":
[1,2,3,4.5] },
  "peptide_measurements": { "pepa1a1":
  {
    "proforma": "SLTNDWEDHLAVK",
    "mods": "free text",
    "rt_target": 853.78,
    "xics": [
      {
        "mz": 764.3755373,
        "charge": 2,
        "isotope": 0,
        "rank": 1,
        "th_ratio": 0.568912,
        "quality": "a",
        "label": "light",
        "trace": {
          "x": [851.78, 852.78, 853.78, 854.78],
          "y": [851.78, 852.78, 853.78, 854.78]
```

```

    },
    "peak": {
      "area": 450245623,
      "max_intensity": 2345.456,
      "rt": [851.78, 852.78, 853.78],
      "aligned_rt": [851.78, 852.78, 853.78]
    }
  },
  {
    "mz": 764.8769,
    "charge": 2,
    "isotope": 1,
    "quality": "a",
    "trace": {
      "x": [851.78, 852.78, 853.78, 854.78],
      "y": [851.78, 852.78, 853.78, 854.78]
    },
    "peak": {
      "area": 550245623,
    }
  }
]
},
"pepala2": {
},
}
}
}

```

output file format

```

{"quantification_data" : [
  {
    "quantify_id": "q1",
    "group_id": "g1",
    "first_pass": {"msruna1": "qr_data_block", "msruna2": "qr_data_block"},
    "second_pass": {"msruna1": "qr_data_block", "msruna2": "qr_data_block"}
  }
]
}

```

MSrun retention time alignment

MS run retention time function

```

std::shared_ptr<pappso::MsRunRetentionTime<QString>> &
mcql::MsRunPeptideList::buildMsRunRetentionTimeSp(const mcql::AlignmentMethodSp
&alignment_method)

```

Peak quality code

aa best quality : many MS2 fragmentation event, only one peak directly detected

zaa same as aa, but this charge state was not directly observed in MS2 fragmentation events in this MSrun

- a** good quality, single MS2 fragmentation event, one peak detected
- za** same as a, but this charge state was not directly observed in MS2 fragmentation events in this MSrun
- ab** many MS2 fragmentation event, but more than one peak detected, the greater peak (area) is chosen, it is obviously fragmented... perhaps a hint to check for peak detection parameters
- zab** same as ab, but this charge state was not directly observed in MS2 fragmentation events in this MSrun
- b** peak obtained by “match between run” on the mean aligned observed retention times in MS2 fragmentation events **and** also matching with the retention time given by other detected and quantified MS1 apex peaks in other MS runs
- c** peak obtained by “match between run” only on the mean of aligned observed retention times in MS2 fragmentation events
- d** peak obtained by “match between run” only matching with the retention time given by other detected and quantified MS1 apex peaks in other MS runs
- missed** no peak detected, no quantification

Match between run difference with legacy

MassChroQ

The match between run process behaviour is slightly different between legacy MassChroQ and MassChroQlite. The new process is more conservative as it will not try to find different peptide charge state if the peptide was observed for an MSrun in an other charge state. This lead to a little bit less peptide/charge quantifications but more reliable results. Also if a peptide was observed in MS2 but never quantified in an MSrun in MS1, this one will not be searched in the match between run process using inferred retention by alignment.