

Analyse correlation among precursors

Functional Genomics Center Zurich

15 September, 2018

```
## [1] TRUE
```

QValue Summaries

```
## [1] "precursor_Id" "srm_QValueMin" "srm_QValueNR"

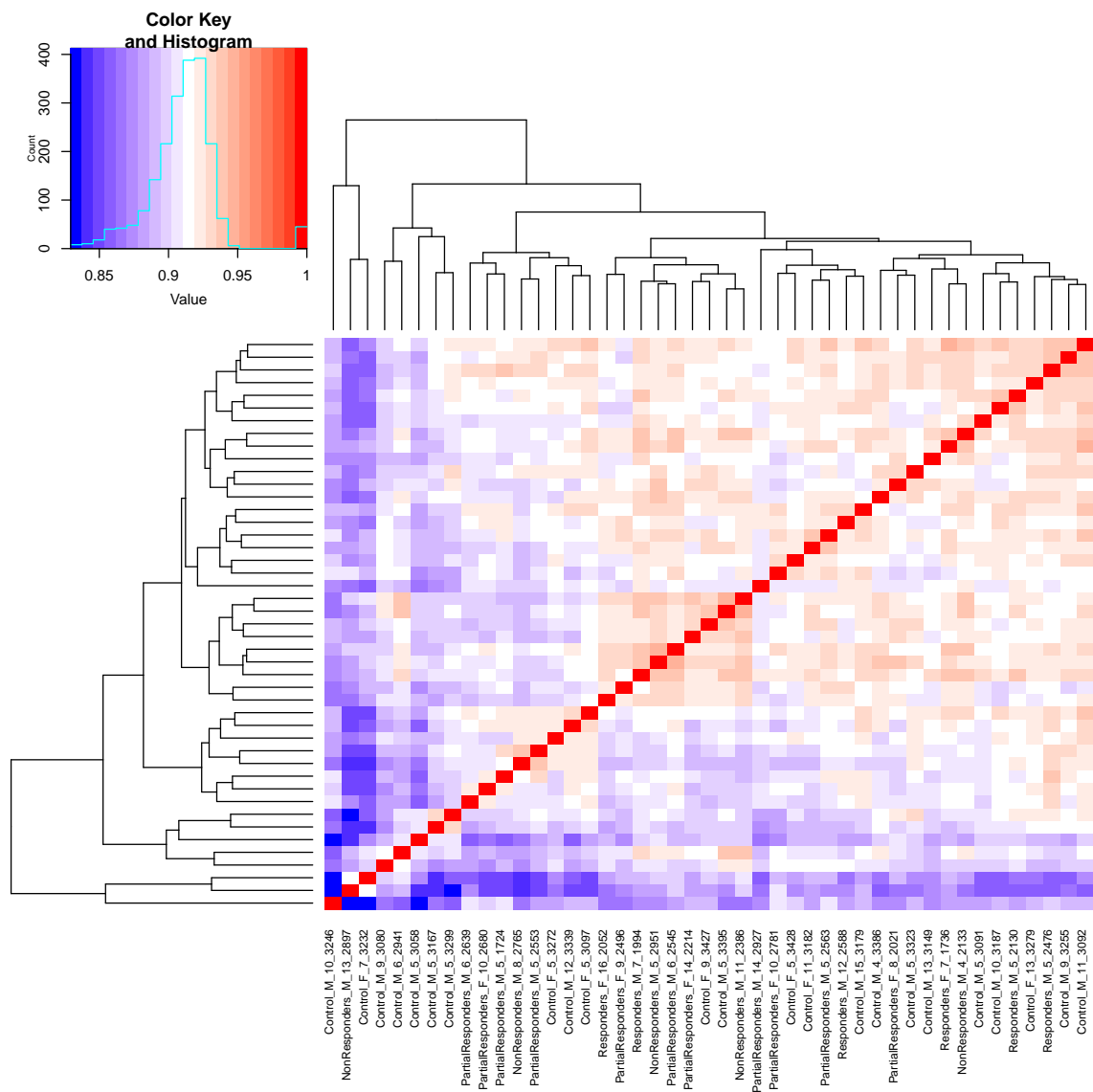
## # A tibble: 16,776 x 7
## # Groups:   coding, protein_Id, precursor_Id [16,776]
##   coding protein_Id precursor_Id Isotope.Label nrReplicates nrNAs
##   <chr>   <chr>      <chr>      <chr>          <int> <int>
## 1 Control P02649    _SELEEQLTPVAEETRAR~ Light          23    23
## 2 Control P04264    _NKLNDLEDALQQA__3 Light          23    23
## 3 Control Q96KN2    _VFQ[+1]YIDLHQDEFV~ Light          23    23
## 4 Control P00450    _DIASGLIGPLIIC[+57~ Light          23    22
## 5 Control P00736    _QRPPDLDTSSN[+1]AV~ Light          23    22
## 6 Control P01023    _NEDSLVFVQTDK__3 Light          23    22
## 7 Control P01024    _QLANGVDYISK__3 Light          23    22
## 8 Control P01024    _VELLHNPAFC[+57]SL~ Light          23    22
## 9 Control P01031    _VDDGVASFVNLPSGVT~ Light          23    22
## 10 Control P01877   _QEPSQGTTTYAVTSILR~ Light          23    22
## # ... with 16,766 more rows, and 1 more variable: meanArea <dbl>
```

Filter by NA's

```
## # A tibble: 1 x 5
##   Isotope.Label protein_Id peptide_Id modPeptide_Id precursor_Id
##   <chr>          <int>      <int>          <int>          <int>
## 1 Light          245      2315          3380          4194

## # A tibble: 1 x 5
##   Isotope.Label protein_Id peptide_Id modPeptide_Id precursor_Id
##   <chr>          <int>      <int>          <int>          <int>
## 1 Light          245      2208          3221          3979

## # A tibble: 1 x 5
##   Isotope.Label protein_Id peptide_Id modPeptide_Id precursor_Id
##   <chr>          <int>      <int>          <int>          <int>
## 1 Light          237      2058          2999          3671
```



Remove single hit wonders

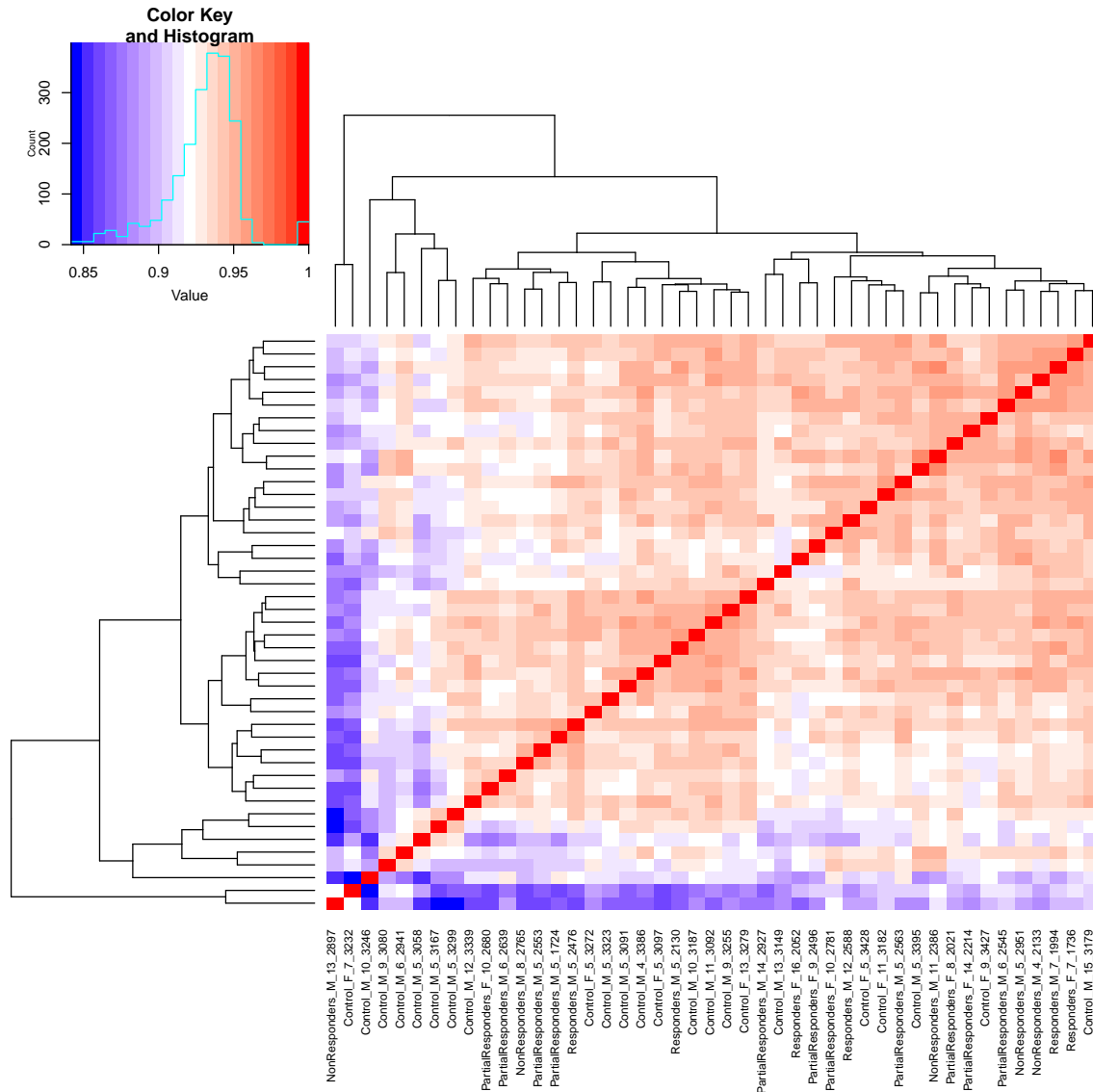
```
## # A tibble: 1 x 5
##   Isotope.Label protein_Id peptide_Id modPeptide_Id precursor_Id
##   <chr>           <int>      <int>      <int>      <int>
## 1 Light             184       2005       2927       3583
```

Correlation Filtering

```
## # A tibble: 1 x 5
```

```
## Isotope.Label protein_Id peptide_Id modPeptide_Id precursor_Id
## <chr> <int> <int> <int> <int>
## 1 Light 184 2005 2927 3583

## # A tibble: 1 x 5
## Isotope.Label protein_Id peptide_Id modPeptide_Id precursor_Id
## <chr> <int> <int> <int> <int>
## 1 Light 145 1455 2138 2619
```



Rank precursors by Intensity

```
## # A tibble: 2,619 x 3
## protein_Id precursor_Id srm_NrNotNAs
## <chr> <chr> <int>
```

```

## 1 AOA075B6P5;P01615 _DIVM[+16]TQSPLSLPVTGPGEASISC[+57]R__3 45
## 2 AOA075B6P5;P01615 _DIVMTQSPLSLPVTGPGEASISC[+57]R__3 41
## 3 AOA075B6P5;P01615 _PGQSPQLLIYLGSN[+1]R__2 27
## 4 AOA075B6P5;P01615 _PGQSPQLLIYLGSNR__2 45
## 5 AOA075B6P5;P01615 _PGQSPQLLIYLGSNR__3 45
## 6 AOA075B6P5;P01615 _SSQSLLHSNGYNYLDWYLQK__3 45
## 7 AOA075B6S2 _FSGSGSGTDFTLK__2 45
## 8 AOA075B6S2 _TYLYWYLQK__2 22
## 9 AOA087WSY6;P01624 _ASQSVSSN[+1]LAWYQQK__2 45
## 10 AOA087WSY6;P01624 _ASQSVSSNLAWYQQK__2 45
## # ... with 2,609 more rows

## [1] "log2_FG.Quantity"

## # A tibble: 2,619 x 9
##   protein_Id precursor_Id srm_QValueMin srm_QValueNR srm_NrNotNAs
##   <chr> <chr> <dbl> <int> <int>
## 1 AOA075B6P5;~ _DIVM[+16]TQSPLSL~ 2.50e-21 45 45
## 2 AOA075B6P5;~ _DIVMTQSPLSLPVTGP~ 6.93e-17 41 41
## 3 AOA075B6P5;~ _PGQSPQLLIYLGSN[+~ 2.86e- 4 27 27
## 4 AOA075B6P5;~ _PGQSPQLLIYLGSNR_~ 1.35e-29 45 45
## 5 AOA075B6P5;~ _PGQSPQLLIYLGSNR_~ 1.19e-63 45 45
## 6 AOA075B6P5;~ _SSQSLLHSNGYNYLDW~ 3.89e-30 45 45
## 7 AOA075B6S2 _FSGSGSGTDFTLK__2 3.15e-64 45 45
## 8 AOA075B6S2 _TYLYWYLQK__2 4.06e- 8 22 22
## 9 AOA087WSY6;~ _ASQSVSSN[+1]LAWY~ 1.27e-20 45 45
## 10 AOA087WSY6;~ _ASQSVSSNLAWYQQK_~ 1.84e-54 45 45
## # ... with 2,609 more rows, and 4 more variables: srm_NrNotNARank <int>,
## # srm_decorelated <lgl>, srm_meanInt <dbl>, srm_meanIntRank <int>

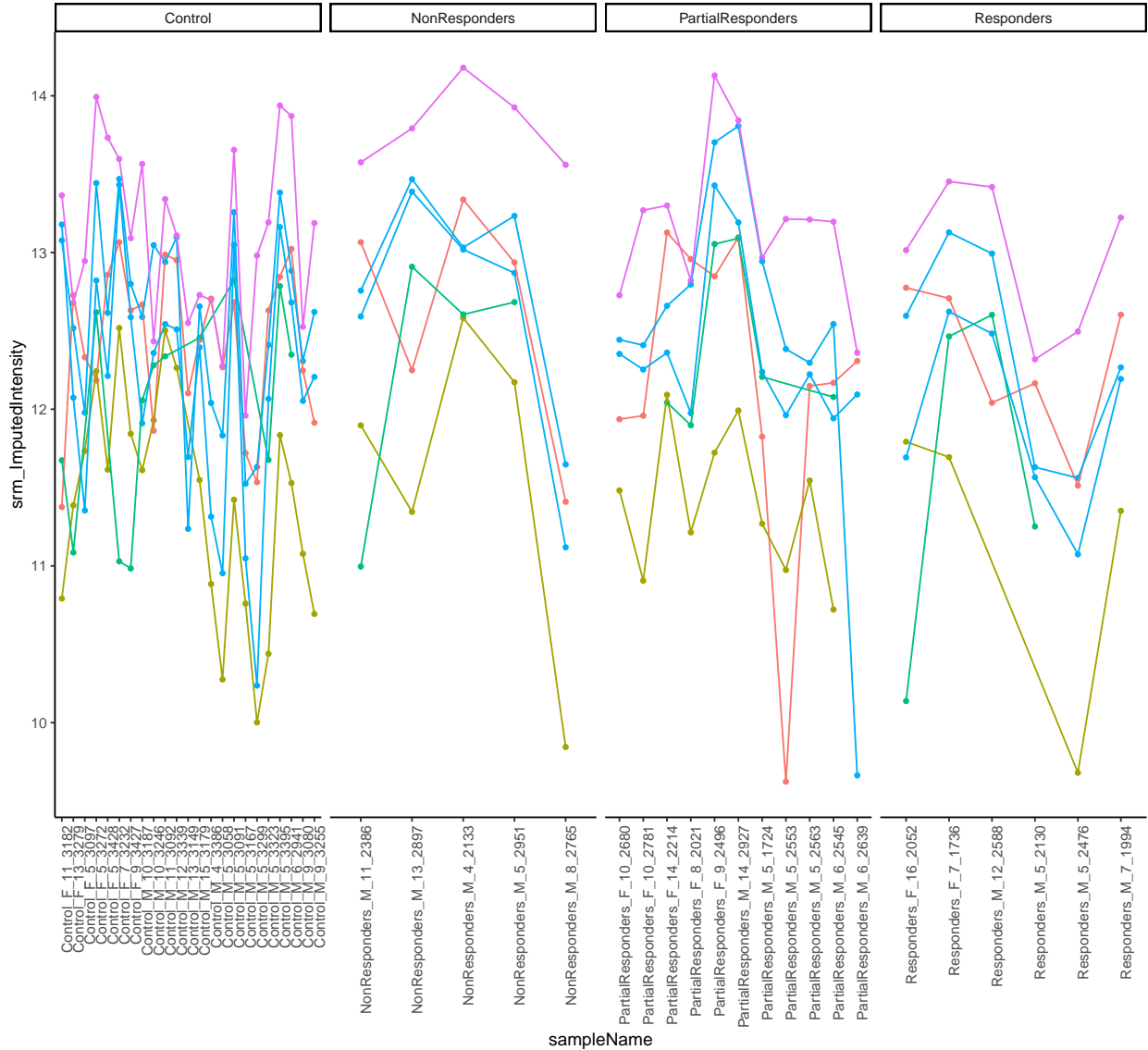
## # A tibble: 110,712 x 10
##   protein_Id precursor_Id srm_QValueMin srm_QValueNR srm_NrNotNAs
##   <chr> <chr> <dbl> <int> <int>
## 1 AOA075B6P5;~ _DIVM[+16]TQSPLSL~ 2.50e-21 45 45
## 2 AOA075B6P5;~ _DIVM[+16]TQSPLSL~ 2.50e-21 45 45
## 3 AOA075B6P5;~ _DIVM[+16]TQSPLSL~ 2.50e-21 45 45
## 4 AOA075B6P5;~ _DIVM[+16]TQSPLSL~ 2.50e-21 45 45
## 5 AOA075B6P5;~ _DIVM[+16]TQSPLSL~ 2.50e-21 45 45
## 6 AOA075B6P5;~ _DIVM[+16]TQSPLSL~ 2.50e-21 45 45
## 7 AOA075B6P5;~ _DIVM[+16]TQSPLSL~ 2.50e-21 45 45
## 8 AOA075B6P5;~ _DIVM[+16]TQSPLSL~ 2.50e-21 45 45
## 9 AOA075B6P5;~ _DIVM[+16]TQSPLSL~ 2.50e-21 45 45
## 10 AOA075B6P5;~ _DIVM[+16]TQSPLSL~ 2.50e-21 45 45
## # ... with 110,702 more rows, and 5 more variables: srm_NrNotNARank <int>,
## # srm_decorelated <lgl>, srm_meanInt <dbl>, srm_meanIntRank <int>,
## # srm_ImputedIntensity <dbl>

```

Aggregation

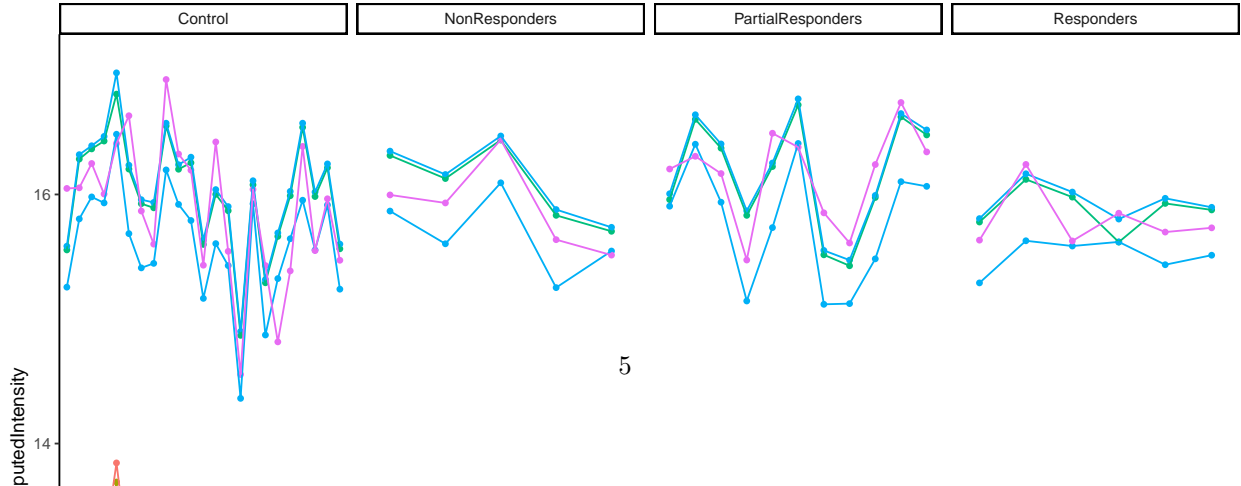
A0A075B6P5;P01615

dPeptide_Id _DIVM[+16]TQSPLSLPVTPGEPASISC[+57]R_ _DIVMTQSPLSLPVTGPGEASISC[+57]R_ _PGQSPQLLIYLGSN[+1]R_ _PGQSPQLLIYLG

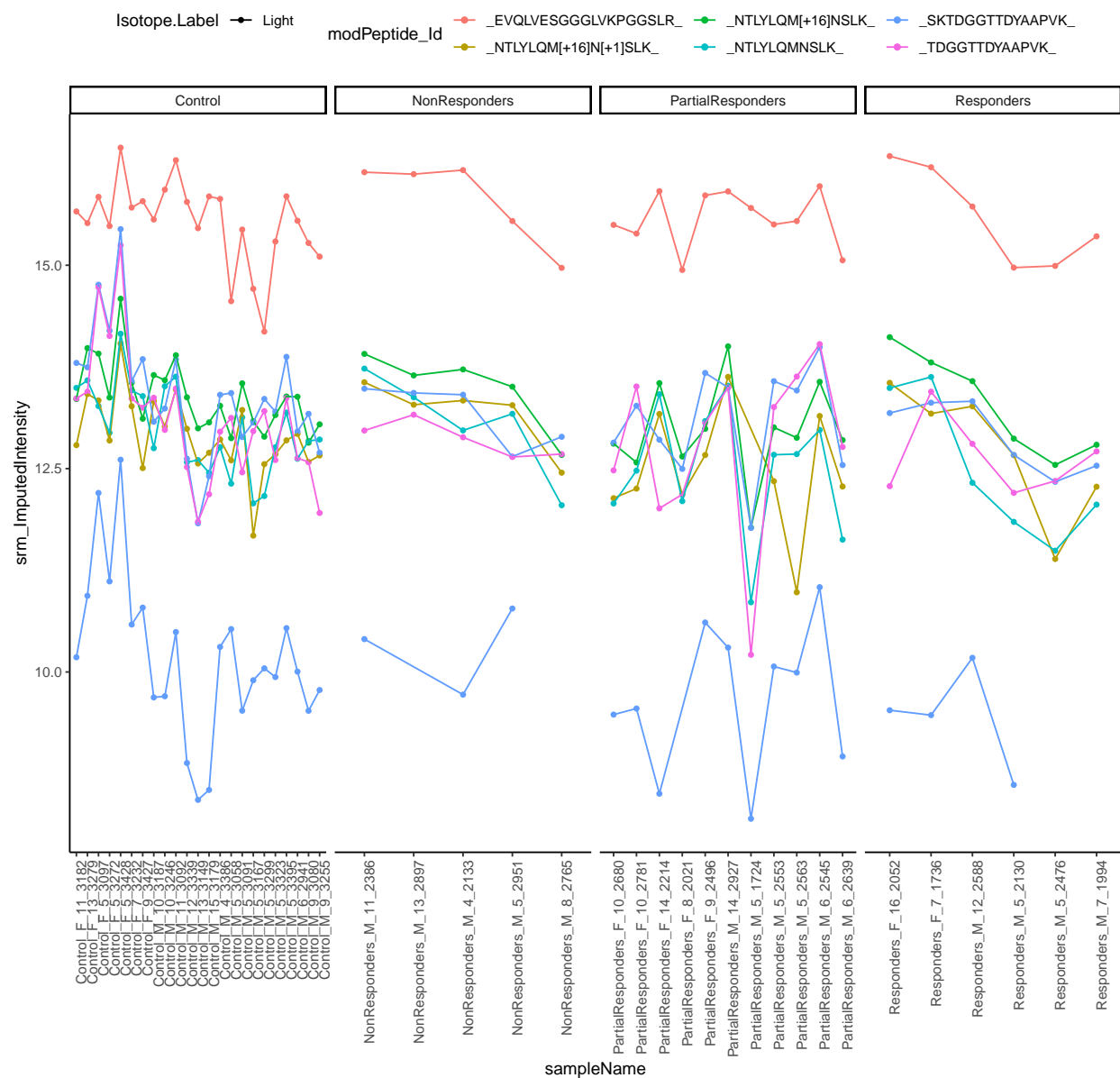


A0A087WSY6;P01624

modPeptide_Id _ASQSVSSN[+1]LAWYQQK_ _ASQSVSSNLAWYQQK_ _EIVM[+16]TQ[+1]SPATLSVSPGER_ _EIVM[+16]TQSPATLSVSPGER_

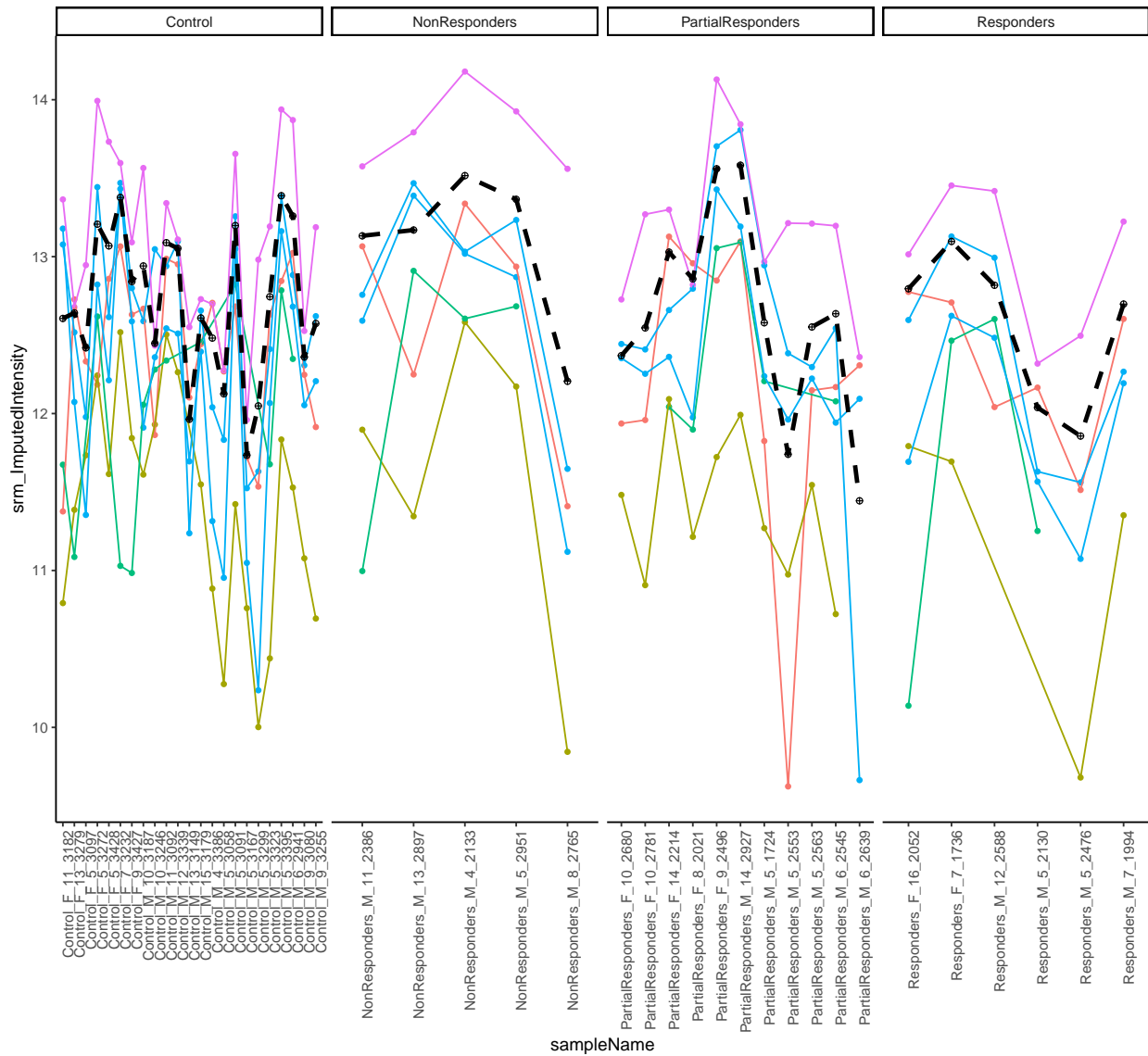


A0A0B4J1V0



A0A075B6P5;P01615

dPeptide_Id _DIVM[+16]TQSPSLPVTGPGEASISC[+57]R_ _DIVMTQSPSLPVTGPGEASISC[+57]R_ _PGQSPQLLIYLGSN[+1]R_ _PGQSPQLLIYLG



Combine all in one function.