**I. PG Extractions for UPLC** - CTCole/Revision E

Procedure summarized from “Phenolic Glycoside Analysis on the Waters Acquity UPLC) by KFR & RLL 11-7-12. For use with “PG processing data sheet”.

**Prep:**

a. get ice for sonicator (add just before sonicating)

b. Turn on cfg to pre-chill (pre-set #1 has 4̊C)

Remove from freezer to come near room temp ca. ½ hr.:

c. PG standards (“Calibration standards): salicortin, tremulacin and tremuloidin are pre-weighed, in big white poly bottle: in dessicator, freezer #3, or in scintillation vials– $20,000 worth in the bag. Salicin is stored RT.

d. Controls (“check standards”): small green bottle: samples used in each run for known results: freezer #3, top shelf, right. Usually do 2 extracts from this each time, processing along with experimental samples. **Record mass** of each.

Remove from freezer just before use:

e. MeOH spiked with 0.01 mg/mL β-resorcylic acid, a process standard (good solubility, size, retention time on UPLC, not found in plants)

ORGANIZATION: Note this uses multiple mcfg tubes etc. for each sample. To avoid labeling all of them, *keep your samples arranged in the same order through each step of the process after the centrifugation*.

**Pipetting**: Tips “wet” with MeOH differently than with water, and accurate measurements *require conditioning the tip* with the solution to be transferred prior to transferring the actual aliquot. Hence each sample is drawn into the tip, ejected, and then drawn back in to transfer. Also, *time must be allowed for the solution to drain down the inner surface of the tip*, and then be blown out, typically 2-6 times.

**Mixing Standards**

**Internal Standard:** “Internal standard” is d6-salicylic acid, 1.00 mg/mL in MeOH. Make enough to match the number of samples for which your calibration standards (PG standards) will be used– not to be stored more than ~2 weeks in freezer. Example: for 638 samples; planning to have some extra, using 30 ug/ sample, aim for ca. 21 mg d6-salicylic acid. Actual weighed amount was 20.45 mg. Add to 20.5 mL MeOH (measured in 50 mL grad cyl), flushing the mcfg tube with MeOH to rinse all the powder out. Cover with Parafilm & shake well to mix. Aliquot into O-ring screw cap mcfg tubes, cover with Parafilm, store at -20 ̊C.

**PG standards (“Analytical Standards for Calibration”):** For *P. tremuloides* we use 4 PG standards. Salicin is obtained commercially, while others (salicortin, tremulacin, tremuloidin) are purified here; lots differ in quality, and degrade into salicin. Stocks of each are made at 1.0 mg/mL in MeOH, but used at different ranges of concentration. See Tables 1& 2, below. Making a large volume of standards and running separately, rather than mixed, greatly reduces batch-to-batch variation.

**Check standards (“controls” or “Aspen lab standards”):** Samples of *P. tremuloides* used for each run since they produce known results. Process 2 extracts each time along with the exptl. samples.

**Storage:**

**PG standards** can be stored for several weeks if sealed tightly and stored at -20 ̊C. Diluted standards in HPLC vials, once punctured, must be removed promptly from the machine, wrapped in parafilm, and stored at -20 ̊C; this is good for about a week. **Extracts** can be stored up to 6 days if the tops are wrapped in Parafilm, and they are stored at -20 ̊C.

**Procedure:**

**A. Prepare PG standards** (“Analytical Standards for Calibration”)

|  |  |  |
| --- | --- | --- |
| 1. Weigh PG standards on micro-balance. Description is for each standard.  2. Use top left button on micro-balance to set to 2 decimal places  3. Tare a tin foil GC “weigh boat”.  4. Handle tin with forceps, place tube on weigh pan & close the door. Zero the balance.  5. Use micro-spatula to add compound. Close door for readings.  6. Dot in upper left corner of balance display disappears when reading is stable. Don’t lean on bench. Aim for 1-3 mg.  7. Dissolve in MeOH (no β-resorcylic acid), adding enough to produce 1.0 mg/mL.  8. Vortex ~ 10 min to fully dissolve.  9. Dilute per Tables below.  10. Filter through a vacuum manifold (as described for extracts, below).  11. store at -20 ̊C.  12. For use, aliquot 20 uL into HPLC tubes. |  | 1. Make these before starting an extraction.  For *P. tremuloides* use salicortin, tremulacin, and tremuloidin (pre-weighed mcfg tubes in freezer 3; brought to room temp before weighing) and salicin (Sigma S-0625 on R.T. shelf).  CONTAINER: For < 2 mL, use screw-cap mcfg tubes with O-ring tops. For 2 - 5 mL use glass dram vials with Teflon-lined caps.  2. Aluminum foil on the lab bench next to the micro-balance also helps reduce static.  3. Glassine can accumulate too much static for stable readings.  6. For Table 1A&B, aim for 1-3 mg. For Table 2, aim for 2 mg (Salicin, Tremuloidin) or 7 mg (Salicortin, Tremulacin).  7. Calculate amount of MeOH to add & record. Don’t put pipettor into bottle: pour small volume into a beaker & pipet from that.  8. Bundle containers with a heavy rubber band. Use 2-4 more rubber bands to hold the bundle to a flat-top vortexer.  10. Large volumes of Table 2 have to be filtered in aliquots < 1 mL.  12. Less might work but hasn’t been tested. |

**Spacer page**

**B. PG Extraction- Samples & Aspen Stds.**

|  |  |  |
| --- | --- | --- |
| 1. Place PG sample mcfg tubes into circular floating test tube racks, open.  2. Add 1.5 mL of ice-cold, spiked MeOH to each mcfg tube.  3. Sonicate 15 minutes in ice bath.  4. Rinse the MeOH dispenser so it doesn’t dry & clog.  5. Centrifuge 3000 rpm for 10 min at <10 C.  6. Set up vacuum manifold. Make sure a spacer plate is in the bottom, then a 96-well receiver plate. 96-well filter plate goes on the cover.  7. Pre-rinse filter with 200 uL MeOH in each well. Can leave vacuum on after MeOH goes through to dry the filter.  8. Set array of new mcfg tubes in rack, one for each sample. |  | 1. Drawer #18. Place hinges of inner row towards center so they’re easier to open. Mcfg tubes labeled with serial numbers, keyed to Tree ID (block/row/position).  2. Spiked MeOH from freezer (MeOH expands substantially from -20 to +20 ̊C. For consistency, we keep the spiked MeOH at -20 ̊C until just before use. Use bottle-top dispenser from cabinet D3. Adjust pickup tube so it sits just above bottom of bottle. First draws to prime & purge air, into a small grad. cyl. Lift pump to the stop, allow to drop slowly to fill tubes. If bottle is empty it gets tippy. Mix by covering rack with big weigh boat & inverting.  3. When ice is added, sonicator should be filled to top row of holes. Set temp to 0, timer to 15 mins.  4. Rinse with unspiked MeOH, then d.i. H2O.  5. Pre-setting #1 has these values. Push the lid down to seal; listen for the compressor to start. Should be < 15 ̊C & declining when starting.  6. Stored in cabinet above UPLC. Round bottom receiver plates work best.  7. This only needs to be done before the first set of samples run through the filter. Use 8-channel pipettor, filling from trough with MeOH. Filters can be damp, need not be dry. 96-  well filters should be marked to indicate which wells have been used. Filters used for extracts turn green, but those used for PG standards don’t change color.  8. Use 96-tube racks; use every other row (A, C, E, G) |

|  |  |  |
| --- | --- | --- |
| 9. Mix 150 uL of MeOH and 30 uL of “internal standard” per sample (inc. Asp. Stds, plus a bit extra for loss); aliquot 180 uL to each tube, allowing drain-down inside tip & blowing out 2-3 times.  10. Transfer 150 uL of each centrifuged extract to these new mcfg tubes. Mix.  11. Transfer diluted extract to filter wells and turn on vacuum for 30 - 60 sec. Tap the filter sharply to dislodge any suspended droplets.  12. Transfer 200 uL of each filtered extract into a HPLC tube. Cap.  13. Remove filter manifold from vacuum manifold.  14. Remove receiver plate, mark used wells, cover with plate cover, and store.  15. Load 48-well rack of HPLC vials into UPLC autosampler, or store. |  | 9. Dispense using a 200 uL pipettor for accuracy. Move mcfg tubes from round cfg racks to a rectangular rack, laid out identical to array of new mcfg tubes; this will be used to set up UPLC queue. Dilution produces concentrations low enough to work in the UPLC, must be accounted for when calculating PG concentrations. “Internal standard” is d6-salicylic acid.  10. *This is the most critical pipetting step*. “Condition” each tip before transferring aliquot, allow drain-down, and blow-out 5-6 times. When rack of diluted samples is filled, cover rack, hold tightly, and shake well.  11. While entire volume goes into filter, missing a few uL will not make a difference, since concentration is set. “Tap sharply” does not mean “pound on”.  12. Have HPLC tubes in blue, 48-well racks. Keep in same order as tubes. Label first and last, but remainder do not need labeling. Include a MeOH blank with PG Standards (UPLC will be set to run this tube first).  13. Avoid inter-mixing with residual drips on bottom of filter: lift manifold cover very carefully, sliding one hand in to shroud receiver plate from anything that might drip from above. 96-well filter can be stored: do not place the filter face down on a lab bench to avoid contamination. Set it down on a 96-well plate if necessary  14. Store in a clean Zip-Lock bag with its receiver plate (hold together with rubber band). Make sure used wells are clearly marked (especially those used to filter colorless PG standards).  15. Filtered extracts can be stored in plastic HPLC vials < 6 days at -20 ̊C, or 2 days at 4̊C, provided that the septa have not been punctured (or if punctured, that they are sealed with Parafilm). |

**Table 1A: Analytical PG Standards for Calibration (“Calibration standards”) preparation**

In parens: amounts when using 200 uL pipettor, more accurate than a 1 mL pipettor. *Pre-condition each tip with MeOH* prior to measuring each compound, then start with highest volumes and work to smallest. Table 1A is for **Salicortin & Tremulacin**, which have higher concentrations in aspens.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Final concn.  of Scort or Tcin standard  (mg/mL) | **Add First**  Add \_\_\_ uL of MeOH | **Add Second**  Add \_ \_ uL of Internal Standard (1 mg/mL) | **Add Third**  Add \_\_\_ uL of 1.0 mg/mL stock | Total volume  (uL) |
| 1.0 | 0 | 30 | 300 (200 + 100) | 330 |
| 0.75 | 75 | 30 | 225 (200 + 25) | 330 |
| 0.375 | 187.5 | 30 | 112.5 | 330 |
| 0.15 | 255 (200 + 55) | 30 | 45 | 330 |
| 0.03 | 291 (200 + 91) | 30 | 9 | 330 |
| 0.015 | 295.5 (200 + 5.5) | 30 | 4.5 | 330 |

See introductory notes for preparation of Internal Standard; this can be stored frozen for weeks.

**Table 1B: Salicin & Tremuloidin standard for calibration**

Much lower concentrations are found in *Populus* so the 1.0 mg/mL is not needed.

Dilute stock 1.0 mg/mL Salicin or Tremuloidin with MeOH and internal standard per the table

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Final concn.  of Scin or Tdin standard  (mg/mL) | **Add First**  Add \_\_ uL of MeOH | **Add Second**  Add \_\_ uL of Internal Standard  (1 mg/mL) | **Add Third**  Add \_ \_ uL of 1.0 mg/mL stock | Total volume  (uL) |
| 0.375 | 187.5 | 30 | 112.5 | 330 |
| 0.15 | 255 (200 + 55) | 30 | 45 | 330 |
| 0.03 | 291 (200 + 91) | 30 | 9 | 330 |
| 0.015 | 295.5 (200 + 5.5) | 30 | 4.5 | 330 |

—> ***FOR LARGE AMOUNTS (3 mL), SEE TABLE 2A & B***

**Table 2: Analytical PG Standards for Calibration (“Calibration standards”) preparation**

–> **LARGE SCALE VERSION (3.3 mL in 1 dram vials with PTFE-lined caps)**

*Pre-condition each tip with MeOH* prior to measuring each compound, then start with highest volumes and work to smallest. Table 2A is for **Salicortin & Tremulacin**, which have higher concentrations in aspens. (in parens: mL to add using 1 mL pipettor).

**Table 2A:** Need > 7 mg Salicortin, Tremulacin into MeOH for stock.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Final concn.  of Scort or Tcin standard  (mg/mL) | **Add First**  Add \_\_\_ uL of MeOH | **Add Second**  Add \_ \_ uL of Internal Standard (1 mg/mL) | **Add Third**  Add \_\_\_ uL of 1.0 mg/mL stock | Total volume  (uL) |
| 1.0 | 0 | 300 | 3000 (1 + 1 + 1) | 3300 |
| 0.75 | 750 | 300 | 2250 (1 + 1 + 250) | 3300 |
| 0.375 | 1875 (1 + 0.875) | 300 | 1125 (1 + 0.125) | 3300 |
| 0.15 | 2550 (1 + 1 + 0.550) | 300 | 450 | 3300 |
| 0.03 | 2910 (1 + 1 + 0.910) | 300 | 90 | 3300 |
| 0.015 | 2955 (1 + 1 + 0.955) | 300 | 45 | 3300 |
| *TOTAL* |  | *1800 uL (3600 for 2 PGs)* | *6960 uL* |  |

See introductory notes for preparation of Internal Standard; this can be stored frozen for weeks.

**Table 2B: Salicin & Tremuloidin standard for calibration**

Much lower concentrations are found in *Populus* so the 1.0 mg/mL is not needed.

Dilute stock 1.0 mg/mL Salicin or Tremuloidin with MeOH and internal standard per the table.

Need > 1.7 mg Salicin, Tremuloidin into MeOH for stock

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Final concn.  of Scin or Tdin standard  (mg/mL) | **Add First**  Add \_\_ uL of MeOH | **Add Second**  Add \_\_ uL of Internal Standard  (1 mg/mL) | **Add Third**  Add \_ \_ uL of 1.0 mg/mL stock | Total volume  (uL) |
| 0.375 | 1875 (1 + 0.875) | 300 | 1125 (1 + 0.125) | 3300 |
| 0.15 | 2550 (1 + 1 + 0.550) | 300 | 450 | 3300 |
| 0.03 | 2910 (1 + 1 + 0.910) | 300 | 90 | 3300 |
| 0.015 | 2955 (1 + 1 + 0.955) | 300 | 45 | 3300 |
| *TOTAL* |  | *1200 uL (2400 for 2 PGs)* | *1710 uL* |  |

spacer page

PG Processing Data Sheet Date:

**PG Standards:**

Date of 1.0 mg/mL stock used: Lot (record when dissolving):

Salicin \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Salicortin \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Tremulacin \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Tremuloidin \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Date PG Standard Dilution Set was made & original Batch it was first used with:

Date: Batch:

**Aspen Control Standards**

Asp1 mg Sample ID:

Asp2 mg Sample ID:

Asp3 mg Sample ID:

**Sample Serial Numbers:**

**n =**

**Data Processed in QuanLynx Batch named:**

**Internal Standard mix**

For 78 samples + 2 Aspen Stds: 13.5 mL MeOH + 2.70 mL internal std. (1.0 ug/uL d6-salicylic acid in MeOH). Note if other volumes used.

**Verify WisAsp column set is installed:**

**Verify Mobile Phase & Wash bottles for the run (A @ 10 mL, B @ 4.4 mL:**

**Verify Aspen Std file (line) names increment (...A, ....B, etc.):**

**After entering MassLynx queue:**

Beginning row in queue: Ending row: n rows:

**UPLC run times** (@26.6 min)

Beginning date: Time: Est. Duration:

Est. completion date: Time:

QuanLynx PG Output

**Calibrations**

Internal Std: check consistency & range (check and/or comments):

PG Stds: R2 values for calibration curve Calib. curve coeffs for x, x2:

Salicin:

Salicortin:

Tremulacin:

Tremuloidin:

Do any PG standards need peak integration adjustments?

Blanks: check peak areas are low (check and/or comments):

Data review: check each compound, for each sample, for appropriate peak integration (correct peak, appropriate ends of tails) (check and/or comments):

Data stored as (filenames):

**NOTE ANY STRANGE-LOOKING SAMPLES THAT NEED RE-DO**

**II. UPLC runs for WisAsp** - CTC

1. Make sure the instrument is ready:

a. check for any red lights

If there are any flashing red lights, the first troubleshooting step is to use the Acquity software to reset that component. For example, if the Binary Solvent Manager has a red light, open the Acquity program, select Binary Solvent Manager > Control > Reset BSM. If that doesn’t work, re-start the Acquity program. If that doesn’t work, find Chris or Mark.

b. check the mobile phase bottles before each run (see Extraction protocol).

c. check the waste bottles

The mobile phase bottle should have room for all the input (goes into Organic Waste), and the water bottle from the N2 generator rarely needs emptying in the winter, more in the summer (goes down the drain).

2. Start Acuity software

Set the Sample manager temps (so they equilibrate while you do everything else):

Column temp = 40̊

Sample temp = 4̊

3. Start the MassLynx software.

4. Set up the project management in MassLynx:

a. Open Project – window opens with a comment about “automatically closed down” – click on “Yes” (this is not a problem)

Fire directory appears – look for C:\\MassLynx\WisAsp2017 folder (etc.)

The Project File is WisAsp 2017 (etc.)

Spreadsheet should appear. Many operations on this spreadsheet work like those of Excel, e.g. inserting rows, using “fill down” or “fill series” in columns, etc. Options can be accessed using right-click.

Check these columns: note a new file is made for each individual sample.

FileName: this will get the sample name, e.g. the sample serial number

Sample ID: this gets the Batch identifier

File Text: a column for notes

MS File: PGS-11 030512SIR-ScanSQ

Inlet File: PGS-11 030512IM

Also note the MS tune file will be PhenGlycs022017

Bottle: format identifies plate and position on plate, e.g. 1:A,1 then 1:A,2 etc. Put the locations of the Calibration Standards and Blank in here manually, then identify the sample that will be in 1:A,1. The remainder will get populated automatically, below.

Injection volume = 2

Sample Type: e.g. Analyte or Standard or Blank

Standard = These are the PG Calibration Standards, run once per Batch

Aspen standards: these are called “Analyte” like the experimental samples, and are re-run roughly every 13 samples; these 2 rows can be copy & insert pasted into the array after the spreadsheet is built.

Click on the Samples tab, enter the number of samples (not including the 2 aspen control standards – you’ll add those later). The spreadsheet is built & populated, e.g. with the bottle numbers (1:A,2 then 1:A,3 etc.).

Then enter File Name. A new file is written for each sample. First should be the Calibration Standards, named this way: ST\_1\_012517 indicates Standard, concentration 1 (of 7), then the date (here January 25, 2017). Run the Blank first, and then run again after each of the PG Calibration Standards, and at the very end of the run. Note that each instance will need a new name, to make a distinct data file.

Aspen lab standards: these are named ASP1\_221\_0308\_A, ASP2\_229\_0308\_A, etc. The digits indicate the weight of the sample (data that will be used later) and then the date (to distinguish aliquots having the same weight; here 8 March). NOTE however that each instance needs a new name so that the files are not over-written, so they should be named something like ASP1\_221\_0308\_A and ASP2\_229\_0308\_A for the first pair of rows, then ASP1\_221\_0308\_B and ASP2\_229\_0308\_B for the second pair, etc.

5. Set up the Acquity program.

a. Binary solvent manager:

A1 = 99%

B1 = 1%

Flow = 0.3 mL/min at start so it does not over-pressure; the program will increase this to 0.5 mL/min when running the samples.

Click on the green check and it starts running. Pressure should be ca. 7000 psi at the 0.3 mL/min; if too high the most likely culprit is the guard column; if too low there’s probably a leak.

b. Sample manager: Verify

Column temp = 40̊

Sample temp = 4̊

c. PDA detector: icon at right edge

d. SQ detector: ion mode = ES- (should usually be negative)

e. Turn on laser using icon on right edge

f. Turn on N2 using API icon on right edge

g. System Status on left edge should be green.

6. To run:

a. Place blue sample plates in the machine. Make sure Plates 1 & 2 are in the correct holders.

b. On the MassLynx spreadsheet, click & drag to highlight the samples to be run. Check to make sure you don’t find any errors in the table.

c. Click on Start Run (blue arrow).

d. Can leave either MassLynx open to note which sample is being processed, or Acquity open to monitor the system pressure.

This program takes about 26.6 minutes to run each sample.

**III. Protocol for Processing UPLC PG data** – C.T. Cole 7 November 2018

UPLC data go through five main steps described below:

A. QuanLynx processing: peak areas to concentrations

B. Restacking (reformatting) data and

C. Reducing to the main four PG compounds

D. Calculating Percent Dry Weight

E. Normalizing data for batch-to-batch variation

F. Ensure you have all the desired data (don’t wait until the very end of a project!)

**A. QuanLynx processing**

1. On the MassLynx table, click & drag to select the samples to quantify. The selection must include the analytical standards of different concentrations (PG Standards). An alternative is to assemble samples and standards that were not run sequentially into a new MassLynx table.

2. Click on Shortcut tab (top), then QuanLynx (left edge), then “Process samples”

3. Verify Method. Click on box next to Method window for a menu of MDB files. Almost always this will be from MS data, e.g. Quantify\_PGs\_11 030512SIR-scan SQ\_CTC-0318.mdb.

4. Check areas of internal standards to verify they are consistent and do not show systematic changes. Record on the back of the PG Processing Data Sheet.

5. For each of the four PGs (salicin, salicortin, tremulacin, and tremuloidin), exclude spurious data points at the bottom of the curve (click and drag to enlarge this portion of the curve). If necessary, also exclude points at the high end of the curve, to maximize the *r*2 value. Ensure that the resulting curve encompasses values found in the sample table.

Since the curve-fitting software uses a quadratic function rather than a saturation curve, excluding higher concentration standards can result in a close fit to low values, but a curve that declines markedly at high values. You want to ensure that no samples have Response values in this upper range. Record the *r*2 and slope values for each PG standard on the back of the PG Processing Data Sheet.

6. Note if any PG standard samples need to have their peak integration areas adjusted manually.

7. Check to ensure that blanks have low values. Record on the PG Processing Data Sheet.

8. For each compound, scroll through each sample to check the peak integration areas and adjust any if necessary (click and drag baseline to adjust).

9. Save the file in the appropriate folder with an appropriate name (e.g. W16\_BATCH37\_1470-2148\_CTC\_19SEPT\_2017.qld). This indicates the project (W = WisAsp), sample year (June 2016 – if August that would be added), Batch ID, serial numbers, who did the QuanLynx processing (CTC), and the date of the QuanLynx processing.

10. Export the file: select “Export complete summary” as a TXT file, using the same name as for the QLD file.

11. Transfer file via a flash drive for further processing.

**B. Restacking (reformatting) and Reducing UPLC data**

All downstream processing (e.g. using R) requires data files that have a row for each sample and a column for each variable, along with a single header row. The QuanLynx output is not arranged in this format, but is organized with blocks of data for each PG compound.

Here is an example of the raw TXT file output from QuanLynx:



Reformatting the data is called “restacking”. While this can be done manually, e.g. in Excel, an R script written by Clay Morrow does this task quickly. (note the first step listed is the same as the last step of part A).

1. Copy TXT file onto flash drive & then onto your computer.

2. Run Clay’s restacker script in R (stored in the lab drive R library – Have to source the code first – i.e. specify path). Here’s an example:

source('C:/LAPTOP BACKUP STUFF/Aspen/CTC R work/Clays R code for PG restacking and selecting CT data with lowest CV\_files/uplc\_restack.R', encoding = 'UTF-8')

uplc.restack(path.in = "C:/LAPTOP BACKUP STUFF/Aspen/UW GWAS/WisAsp PG data/WisAsp PG 2016/UPLC files/UPLC TXT files/UPLC TXT files June 2016 for restacking/W16\_BATCH42\_ Redos JunAug-Comp-Expr\_CTC\_14rev15DEC\_2017.txt", path.out="C:/LAPTOP BACKUP STUFF/Aspen/UW GWAS/WisAsp PG data/WisAsp PG 2016/UPLC output restacked/", multiple=FALSE, wide=TRUE)

The restacker program requires specifying paths for input and output, whether multiple files are to be used or just one file, and whether a wide format is to be used or not (yes).

Here is a restacked file (columns go up through AN):



**C. Reducing the file to the main PG compounds**

The QLD and TXT files record data for many compounds, and to reduce the file to the main four PG compounds of interest, the following steps (**C** and **D**) are handled either in Excel, if there is only one or a few files, or with the “R code for UPLC wrangling MASTER reduced” if there are a bunch of files being processed. The steps below are for Excel, but summarize what the R code does.

1. Open the restacked file in Excl.

2. Reduce the file as follows:

a. Remove the un-named column, Obs, num, and Type columns

b. Remove the columns for compounds except Salicin, Salicortin, Tremuloidin, and Tremulacin (that is the order they appear in, L to R)

c. Remove the rows for Blanks and PG Standards

3. Save the file with the same name but “restacked” changed to “reduced” at the end of the file name.

4. For each compound there are 3 columns with the same name. L to R those 3 are Area, Response, and Concentration, but the column labels say \_NA.

a. For each compound, change the third column from \_NA to an abbreviated name plus Conc: e.g. “ Salicin\_2ion\_NA” becomes ScinConc. Others are ScortConc, TcinConc, and TdinConc

b. Remove the first two columns (they are Area and Response) for each compound.

c. Swap the TcinConc and TdinConc columns so they are in alphabetical order.

d. Add a column for SerialNo. This can be filled using the “Text to Columns” feature of Excel, from data in the Name column.

5. Check the data file to make sure there are data for all of the PG compounds, for each sample. In some cases, data will be missing for one or more PG.

Here is an example of the file reduced (but with columns from the next step, too):



**D. Calculating Percent Dry Weight**

1. Add in a column of sample weights (PGwts). These would have been recorded in a separate data file. Paste these data into the column (or, if many files, use dplyr in R).

2. Add a column for percent dry weight for each of the 4 POGs: ScinDW, ScortDW, TcinDW, TdinDW.

3. Use the sample weights and the Conc data to calculate percent dry weight for each compound.

DW = 1.5 \* 2 \* 1.09 \* 100 \* Conc/PGwt

1.5 = volume

2 = factor for dilution in MeOH

1.09 = correction for 92% extraction efficiency

100 = convert to percent

4. Save the file with these changes.

**E. Normalizing data for batch-to-batch variation**

Running separate PG Standards, across multiple batches, has greatly reduced this variation. Nevertheless, normalizing can take an enormous amount of time. Several sources of variation (“error” in the statistical sense) conspire, and the task is to identify and correct for identifiable systematic and random error.

First, consider how the UPLC absorbance data are quantified. The area for each compound, for each sample, is compared to the area of the internal standard, d6-salicylic acid, which is included with each sample (and each PG standard). The ratio of their areas is the “response factor”, Rf. For each compound, a standard curve is created from the Rf values of different, known concentrations, and used to convert the Rf value for each sample to an equivalent concentration. Sample weight data are then used to convert the concentrations d to percent dry weights.

Now consider sources of error:

a. weighing internal standard used in PG standards (e.g. from static or other problems)

b. pipetting variation of internal standard in PG standards

c. weighing internal standard used in samples

d. pipetting variation of internal standard among samples (likely to be very small if the internal standard is mixed with the MeOH before diluting extracts)

e. weighing PG standards

f. pipetting variation in making Mix A

g. pipetting variation in making the PG Standard dilutions

h. weighing the Aspen Lab Standard Control powder

i. weighing the sample aliquot powder

j. pipetting variation during extraction

k. UPLC instrumental variation, e.g. injection volumes, column temperatures

While careful adherence to the protocol can reduce these sources of error, they can’t be eliminated, and remain large enough to affect the data. Batch-to-batch variation can be revealed by comparing PG concentrations of the Aspen Lab Standard Controls, which should be identical from one batch to another. If concentration of a PG is low in all of the Controls used in a batch, compared to the grand mean across all batches, then we might infer that some source of error would be shared with all the samples of that batch, and we can normalize the values of that batch to the grand mean from all batches. On the other hand, the low value for the Controls might arise from bogus weight data for that Control; in that case, it should not be used to normalize the batch data. How can you tell the difference?

a. Run multiple Aspen Lab Standard Control aliquots in each batch

If all of the controls are low or high, then the variation is probably shared with the samples and the batch should be normalized; if only one control is anomalous, the weight for that control is probably erroneous. In that case, those control data should not be used for normalizing or for calculation of the grand mean.

b. Run each Control multiple times in each batch

Usually the multiple injections of each control have very similar responses, but if they vary systematically, e.g. declining during a batch, then there is probably something going wrong with the batch and it should be re-run.

If values for only one PG standard are anomalous, there probably was an error in weighing that PG standard. In that case, the sample data can be re-processed in QuanLynx by using a different set of PG standards, e.g. from a different run.

If all the controls in a batch are low or high, data for that batch can be normalized by multiplying the batch data by (grand mean/batch control mean). R code for doing this is in the file “R code for correcting PG values to Control means”. The grand mean is taken across many batches, e.g. all of the June 2016 samples.

Here is a portion of the Aspen Standard Controls spreadsheet, using conditional formatting to aid comparison of PG amounts:



**F. Do you have all the data?**

If you have hundreds or thousands of samples, how do you ensure you have all the data, or that there are no duplications?

For the WisAsp and Expression Garden studies, each tree has a unique serial number (SerialNo). In the WisAsp study, these are sequential, starting at 1 through 1568 for the main blocks, and 2001 through 2256 for the border rows. For the Expression Garden, these are 3001 through 3196.

In a spreadsheet of data, such as the PG data, sort the data by SerialNo from low to high. Insert two new columns. The first of these is a Counter, filled sequentially from 1, 2, .... or 2001, 2002.... The second is a Flag, filled with (SerialNo - Counter). This can be given a conditional format (e.g. bold, red) for any value that is not 0, identifying any missing or duplicated numbers.

In the example below, SerialNo 6 is missing data (the tree was dead, and values should be replaced with “NA”):

