Checking discrepancies

Johanna Jantzen

November 18, 2020

Etienne download versus my fulcrum download differences

```
#read in spectral data object

Etienne_data <- read.csv("https://web.fulcrumapp.com/shares/5c5290b884ab3734.csv") # Fulcrum data share

Shan_spectra <- readRDS("./data/all_spectra.rds")

#get metadata object from spectra
metadata_Shan <- meta(Shan_spectra)

#get species list from spectral data
species_Shan <- meta(Shan_spectra, "species", simplify = TRUE)

#get metadata from additional metadata file (downloaded separately from fulcrum)
Johanna_data <- read.csv("./data/metadata/leaf_spectra.csv", stringsAsFactors = FALSE)
```

Checking for differences in my and Etienne's metadata

```
#check which sample IDS are not in my sample IDs
Johanna_missing <- Etienne_data[-which(Etienne_data$sample_id %in% Johanna_data$sample_id),]

Etienne_missing <- Johanna_data[-which(Johanna_data$sample_id %in% Etienne_data$sample_id),]



nrow(Johanna_missing)
```

```
## [1] 461
```

```
nrow(Etienne_missing)
```

```
## [1] 0
```

Mine is missing 461 and Etienne's is missing 0 samples

```
#there are no NA scientific names in Etienne's data
Etienne_data[which(is.na(Etienne_data$scientific_name)),]
```

```
##  [1] fulcrum_id                      created_at
##  [3] updated_at                      created_by
##  [5] updated_by                      system_created_at
##  [7] system_updated_at               version
##  [9] status                          project
## [11] assigned_to                     latitude
## [13] longitude                       geometry
## [15] filter_site                     filter_site_id
## [17] sample                          sample_remarks
## [19] site_id                         sample_id
## [21] scientific_name                 date_measured
## [23] measured_by                     measured_by_other
## [25] spectroradiometer_start_time    event_remarks
## [27] computer                        computer_type
## [29] spectroradiometer_id            manufacturer_short_name
## [31] serial_number                   instrumentation_id
## [33] manufacturer_short_name_sphere  instrumentation_type
## [35] panel_id                        parent_directory
## [37] working_folder                  base_file_name
## [39] leaf_larger_than_port           protocol
## [41] protocol_other                  protocols
## [43] protocol_url                    properties_measured
## [45] leaf_sides_measured             leaf_photos
## [47] leaf_photos_caption             leaf_photos_url
## [49] number_of_measurements          record_is_calculated
## [51] calculated_record_link          quality_leafs
## [53] leaves_without_good_quality     quality_leaves_comments
## [55] deleted_by                      date_deleted
## [57] rejected_by                     date_rejected
## [59] verified_by                     date_verified
## [61] submitted_by                    date_submitted
## [63] approved_by                     date_approved
## [65] published_by                    date_published
## [67] number_of_rejections            spectral_measurements_visibility
## [69] gps_altitude                    gps_horizontal_accuracy
## [71] gps_vertical_accuracy           gps_speed
## [73] gps_course
## <0 rows> (or 0-length row.names)
```

```
#there are no NA scientific name in Johanna's data
Johanna_data[which(is.na(Johanna_data$scientific_name)),]
```

```
##  [1] record_id                 project
##  [3] geometry                  latitude
##  [5] longitude                 altitude
##  [7] sample                    sample_remarks
##  [9] site_id                   sample_id
## [11] scientific_name           date_measured
## [13] measured_by               event_remarks
## [15] working_folder            base_file_name
## [17] properties_measured       leaf_sides_measured
## [19] measurements              number_of_measurements
## [21] quality_leafs             leaves_without_good_quality
## [23] quality_leaves_comments
```

```
## <0 rows> (or 0-length row.names)
```

```
#there are 126 names in Etienne's data
unique(Etienne_data$scientific_name)
```

```
##   [1]
##   [2] Acer pensylvanicum Linnaeus
##   [3] Acer saccharum Marshall
##   [4] Abies lasiocarpa (Hooker) Nuttall
##   [5] Phragmites australis (Cavanilles) Trinius ex Steudel subsp. australis
##   [6] Rhododendron groenlandicum (Oeder) Kron & Judd
##   [7] Abies Miller
##   [8] Plectritis congesta (Lindley) de Candolle
##   [9] Polystichum munitum (Kaulfuss) C. Presl
##  [10] Bromus inermis Leysser
##  [11] Thuja occidentalis Linnaeus
##  [12] Acer saccharinum Linnaeus
##  [13] Abies balsamea (Linnaeus) Miller
##  [14] Alseis blackiana NA
##  [15] Acer rubrum Linnaeus
##  [16] Agonis flexuosa (Willd.) Sweet
##  [17] Fraxinus nigra Marshall
##  [18] Rhus typhina Linnaeus
##  [19] Populus tremuloides Michaux
##  [20] Betula populifolia Marshall
##  [21] Sanicula crassicaulis Poeppig ex de Candolle
##  [22] Prunus pensylvanica Linnaeus f.
##  [23] Quercus macrocarpa Michaux
##  [24] Solidago Linnaeus
##  [25] Carya ovata (Miller) K. Koch
##  [26] Fagus grandifolia Ehrhart
##  [27] Kalmia angustifolia Linnaeus var. angustifolia
##  [28] Eriophorum vaginatum subsp. spissum (Fernald) HultÃ©n
##  [29] Chamaedaphne calyculata (Linnaeus) Moench
##  [30] Kalmia angustifolia Linnaeus
##  [31] Vitis riparia Michaux
##  [32] Rubus idaeus Linnaeus
##  [33] Eriophorum vaginatum Linnaeus subsp. vaginatum
##  [34] Asclepias syriaca Linnaeus
##  [35] Solidago gigantea Aiton
##  [36] Solidago altissima Linnaeus
##  [37] Cornus sericea Linnaeus
##  [38] Phalaris arundinacea Linnaeus
##  [39] Apocynum androsaemifolium Linnaeus
##  [40] Salix interior Rowlee
##  [41] Euthamia graminifolia (Linnaeus) Nuttall
##  [42] Calamagrostis canadensis (Michaux) Palisot de Beauvois
##  [43] Quercus rubra Linnaeus
##  [44] Camassia leichtlinii (Baker) S. Watson
##  [45] Camassia quamash (Pursh) Greene
##  [46] Dactylis glomerata Linnaeus
##  [47] Lomatium utriculatum (Nuttall ex Torrey & A. Gray) J.M. Coulter & Rose
##  [48] Berberis aquifolium Pursh
##  [49] Bromus sterilis Linnaeus
```

```
##  [50] Symphoricarpos albus (Linnaeus) S.F. Blake
##  [51] Festuca idahoensis Elmer
##  [52] Bromus sitchensis var. carinatus (Hooker & Arnott) R.E. Brainerd & Otting
##  [53] Claytonia perfoliata Donn ex Willdenow
##  [54] Oemleria cerasiformis (Torrey & A. Gray ex Hooker & Arnott) J.W. Landon
##  [55] Holodiscus discolor (Pursh) Maximowicz
##  [56] Rosa nutkana C. Presl
##  [57] Quercus garryana Douglas ex Hooker
##  [58] Vicia sativa Linnaeus
##  [59] Poa pratensis Linnaeus
##  [60] Betula alleghaniensis Britton
##  [61] Tilia americana Linnaeus
##  [62] Lathyrus sphaericus Retzius
##  [63] Populus deltoides W. Bartram ex Marshall
##  [64] Crataegus monogyna Jacquin
##  [65] Carpinus caroliniana Walter
##  [66] Sericocarpus rigidus Lindley
##  [67] Acer negundo Linnaeus
##  [68] Cytisus scoparius (Linnaeus) Link
##  [69] Ostrya virginiana (Miller) K. Koch
##  [70] Tsuga canadensis (Linnaeus) CarriÃ¨re
##  [71] Ulmus americana Linnaeus
##  [72] Fraxinus americana Linnaeus
##  [73] Fraxinus pennsylvanica Marshall
##  [74] Pinus strobus Linnaeus
##  [75] Prunus serotina Ehrhart
##  [76] Ulmus rubra Muhlenberg
##  [77] Carya cordiformis (Wangenheim) K. Koch
##  [78] Quercus bicolor Willdenow
##  [79] Quercus alba Linnaeus
##  [80] Juglans cinerea Linnaeus
##  [81] Celtis occidentalis Linnaeus
##  [82] Acer nigrum F. Michaux
##  [83] Pinus resinosa Aiton
##  [84] Quercus Linnaeus
##  [85] Acer platanoides Linnaeus
##  [86] Juglans nigra Linnaeus
##  [87] Aesculus hippocastanum Linnaeus
##  [88] Populus balsamifera Linnaeus
##  [89] Acer spicatum Lamarck
##  [90] Pinus banksiana Lambert
##  [91] Larix laricina (Du Roi) K. Koch
##  [92] Picea mariana (Miller) Britton, Sterns & Poggenburgh
##  [93] Picea rubens Sargent
##  [94] Betula papyrifera Marshall
##  [95] Populus grandidentata Michaux
##  [96] Prunus nigra Aiton
##  [97] Picea abies (Linnaeus) H. Karsten
##  [98] Picea glauca (Moench) Voss
##  [99] Sorbus decora (Sargent) C.K. Schneider
## [100] Lythrum salicaria Linnaeus
## [101] Phragmites australis (Cavanilles) Trinius ex Steudel
## [102] Cirsium arvense (Linnaeus) Scopoli
## [103] Rhamnus cathartica Linnaeus
```

```
## [104] Alnus incana subsp. rugosa (Du Roi) R.T. Clausen
## [105] Salix alba Linnaeus
## [106] Typha angustifolia Linnaeus
## [107] Vicia cracca Linnaeus
## [108] Frangula alnus Miller
## [109] Spiraea alba Du Roi
## [110] Rubus idaeus subsp. strigosus (Michaux) Focke
## [111] Solidago rugosa Miller
## [112] Typha Linnaeus
## [113] Typha latifolia Linnaeus
## [114] Swartzia simplex (Sw.) Spreng.
## [115] Gustavia superba (Kunth) O. Berg
## [116] Trophis racemosa (L.) Urb.
## [117] Inga punctata Willd.
## [118] Anacardium excelsum (Bertero & Balb. ex Kunth) Skeels
## [119] Trichilia tuberculata (Triana & Planch.) C. DC.
## [120] Garcinia madruno (Kunth) Hammel
## [121] Poulsenia armata (Miq.) Standl.
## [122] Heisteria concinna NA
## [123] Protium tenuifolium (Engl.) Engl.
## [124] Virola sebifera Aubl.
## [125] Pinus rigida P. Miller
## [126] Sorbus americana Marshall
## 126 Levels:  ... Vitis riparia Michaux
```

```
#there are 91 names in Johanna's data therefore 35 species names are missing in Johanna's metadata
unique(Johanna_data$scientific_name)
```

```
##  [1] "Acer saccharum Marshall"
##  [2] "Betula populifolia Marshall"
##  [3] "Acer rubrum Linnaeus"
##  [4] ""
##  [5] "Populus tremuloides Michaux"
##  [6] "Kalmia angustifolia Linnaeus var. angustifolia"
##  [7] "Rhododendron groenlandicum (Oeder) Kron & Judd"
##  [8] "Chamaedaphne calyculata (Linnaeus) Moench"
##  [9] "Eriophorum vaginatum subsp. spissum (Fernald) HultÃ©n"
## [10] "Betula papyrifera Marshall"
## [11] "Populus grandidentata Michaux"
## [12] "Fagus grandifolia Ehrhart"
## [13] "Quercus rubra Linnaeus"
## [14] "Eriophorum vaginatum Linnaeus subsp. vaginatum"
## [15] "Kalmia angustifolia Linnaeus"
## [16] "Acer saccharinum Linnaeus"
## [17] "Rubus idaeus Linnaeus"
## [18] "Solidago gigantea Aiton"
## [19] "Solidago altissima Linnaeus"
## [20] "Cornus sericea Linnaeus"
## [21] "Vitis riparia Michaux"
## [22] "Asclepias syriaca Linnaeus"
## [23] "Phragmites australis (Cavanilles) Trinius ex Steudel subsp. australis"
## [24] "Salix interior Rowlee"
## [25] "Euthamia graminifolia (Linnaeus) Nuttall"
## [26] "Apocynum androsaemifolium Linnaeus"
```

```
## [27] "Calamagrostis canadensis (Michaux) Palisot de Beauvois"
## [28] "Phalaris arundinacea Linnaeus"
## [29] "Abies Miller"
## [30] "Agonis flexuosa (Willd.) Sweet"
## [31] "Plectritis congesta (Lindley) de Candolle"
## [32] "Pinus strobus Linnaeus"
## [33] "Tsuga canadensis (Linnaeus) CarriÃ¨re"
## [34] "Solidago rugosa Miller"
## [35] "Betula alleghaniensis Britton"
## [36] "Tilia americana Linnaeus"
## [37] "Bromus inermis Leysser"
## [38] "Populus deltoides W. Bartram ex Marshall"
## [39] "Typha Linnaeus"
## [40] "Carpinus caroliniana Walter"
## [41] "Abies balsamea (Linnaeus) Miller"
## [42] "Rubus idaeus subsp. strigosus (Michaux) Focke"
## [43] "Cirsium arvense (Linnaeus) Scopoli"
## [44] "Alnus incana subsp. rugosa (Du Roi) R.T. Clausen"
## [45] "Phragmites australis (Cavanilles) Trinius ex Steudel"
## [46] "Ulmus americana Linnaeus"
## [47] "Solidago Linnaeus"
## [48] "Fraxinus pennsylvanica Marshall"
## [49] "Fraxinus americana Linnaeus"
## [50] "Quercus macrocarpa Michaux"
## [51] "Acer negundo Linnaeus"
## [52] "Pinus resinosa Aiton"
## [53] "Frangula alnus Miller"
## [54] "Rhamnus cathartica Linnaeus"
## [55] "Spiraea alba Du Roi"
## [56] "Ulmus rubra Muhlenberg"
## [57] "Prunus serotina Ehrhart"
## [58] "Salix alba Linnaeus"
## [59] "Carya ovata (Miller) K. Koch"
## [60] "Carya cordiformis (Wangenheim) K. Koch"
## [61] "Quercus alba Linnaeus"
## [62] "Quercus bicolor Willdenow"
## [63] "Quercus Linnaeus"
## [64] "Juglans cinerea Linnaeus"
## [65] "Rhus typhina Linnaeus"
## [66] "Typha angustifolia Linnaeus"
## [67] "Vicia cracca Linnaeus"
## [68] "Typha latifolia Linnaeus"
## [69] "Lythrum salicaria Linnaeus"
## [70] "Celtis occidentalis Linnaeus"
## [71] "Thuja occidentalis Linnaeus"
## [72] "Fraxinus nigra Marshall"
## [73] "Picea rubens Sargent"
## [74] "Acer pensylvanicum Linnaeus"
## [75] "Populus balsamifera Linnaeus"
## [76] "Picea mariana (Miller) Britton, Sterns & Poggenburgh"
## [77] "Prunus nigra Aiton"
## [78] "Prunus pensylvanica Linnaeus f."
## [79] "Picea glauca (Moench) Voss"
## [80] "Picea abies (Linnaeus) H. Karsten"
```

```
## [81] "Larix laricina (Du Roi) K. Koch"
## [82] "Sorbus decora (Sargent) C.K. Schneider"
## [83] "Acer spicatum Lamarck"
## [84] "Acer nigrum F. Michaux"
## [85] "Pinus banksiana Lambert"
## [86] "Acer platanoides Linnaeus"
## [87] "Ostrya virginiana (Miller) K. Koch"
## [88] "Juglans nigra Linnaeus"
## [89] "Aesculus hippocastanum Linnaeus"
## [90] "Sorbus americana Marshall"
## [91] "Pinus rigida P. Miller"
```

```
#there are samples of 71 species missing from Johanna's metadata that are in Etienne's (maybe not uniqu
unique(Johanna_missing$scientific_name)
```

```
##  [1]
##  [2] Acer pensylvanicum Linnaeus
##  [3] Acer saccharum Marshall
##  [4] Abies lasiocarpa (Hooker) Nuttall
##  [5] Plectritis congesta (Lindley) de Candolle
##  [6] Polystichum munitum (Kaulfuss) C. Presl
##  [7] Phragmites australis (Cavanilles) Trinius ex Steudel subsp. australis
##  [8] Acer saccharinum Linnaeus
##  [9] Alseis blackiana NA
## [10] Acer rubrum Linnaeus
## [11] Agonis flexuosa (Willd.) Sweet
## [12] Sanicula crassicaulis Poeppig ex de Candolle
## [13] Populus tremuloides Michaux
## [14] Betula populifolia Marshall
## [15] Vitis riparia Michaux
## [16] Rubus idaeus Linnaeus
## [17] Solidago gigantea Aiton
## [18] Cornus sericea Linnaeus
## [19] Euthamia graminifolia (Linnaeus) Nuttall
## [20] Camassia leichtlinii (Baker) S. Watson
## [21] Camassia quamash (Pursh) Greene
## [22] Dactylis glomerata Linnaeus
## [23] Lomatium utriculatum (Nuttall ex Torrey & A. Gray) J.M. Coulter & Rose
## [24] Berberis aquifolium Pursh
## [25] Bromus sterilis Linnaeus
## [26] Symphoricarpos albus (Linnaeus) S.F. Blake
## [27] Festuca idahoensis Elmer
## [28] Bromus sitchensis var. carinatus (Hooker & Arnott) R.E. Brainerd & Otting
## [29] Claytonia perfoliata Donn ex Willdenow
## [30] Oemleria cerasiformis (Torrey & A. Gray ex Hooker & Arnott) J.W. Landon
## [31] Holodiscus discolor (Pursh) Maximowicz
## [32] Rosa nutkana C. Presl
## [33] Quercus garryana Douglas ex Hooker
## [34] Vicia sativa Linnaeus
## [35] Poa pratensis Linnaeus
## [36] Lathyrus sphaericus Retzius
## [37] Crataegus monogyna Jacquin
## [38] Sericocarpus rigidus Lindley
## [39] Tilia americana Linnaeus
```

```
## [40] Betula alleghaniensis Britton
## [41] Acer negundo Linnaeus
## [42] Cytisus scoparius (Linnaeus) Link
## [43] Ostrya virginiana (Miller) K. Koch
## [44] Fagus grandifolia Ehrhart
## [45] Populus deltoides W. Bartram ex Marshall
## [46] Carya ovata (Miller) K. Koch
## [47] Carya cordiformis (Wangenheim) K. Koch
## [48] Tsuga canadensis (Linnaeus) CarriÃ¨re
## [49] Quercus alba Linnaeus
## [50] Carpinus caroliniana Walter
## [51] Juglans cinerea Linnaeus
## [52] Ulmus americana Linnaeus
## [53] Pinus strobus Linnaeus
## [54] Fraxinus nigra Marshall
## [55] Ulmus rubra Muhlenberg
## [56] Swartzia simplex (Sw.) Spreng.
## [57] Gustavia superba (Kunth) O. Berg
## [58] Trophis racemosa (L.) Urb.
## [59] Inga punctata Willd.
## [60] Anacardium excelsum (Bertero & Balb. ex Kunth) Skeels
## [61] Trichilia tuberculata (Triana & Planch.) C. DC.
## [62] Garcinia madruno (Kunth) Hammel
## [63] Poulsenia armata (Miq.) Standl.
## [64] Heisteria concinna NA
## [65] Protium tenuifolium (Engl.) Engl.
## [66] Virola sebifera Aubl.
## [67] Quercus macrocarpa Michaux
## [68] Fraxinus pennsylvanica Marshall
## [69] Asclepias syriaca Linnaeus
## [70] Salix interior Rowlee
## [71] Fraxinus americana Linnaeus
## 126 Levels:  ... Vitis riparia Michaux
```

```
#these missing samples include samples from four categories of data
unique(Johanna_missing$status)
```

```
## [1] pending   deleted   verified  submitted
## Levels: deleted pending rejected submitted verified
```

```
#there are samples from 8 projects missing in my metadata that are in Etienne's
unique(Johanna_missing$project)
```

```
## [1] CABO-test                    2018-Hacker-PhD-UBC
## [3] 2018-Boucherville            2019-Blanchard-MSc-UdeM
## [5] 2019-Pardo-MSc-UdeM          SWA-Warren
## [7] 2018-BeauchampRioux-MSc-UdeM 2019-Boucherville
## 13 Levels: 2018-BeauchampRioux-MSc-UdeM ... SWA-Warren
```

```
#which project are the deleted ones missing from? 3: CABO-test, 2018-Hacker-PhD-UBC and 2019-Blanchard-
unique(Johanna_missing$project[which(Johanna_missing$status == "deleted")])
```

```
## [1] CABO-test                2018-Hacker-PhD-UBC     2019-Blanchard-MSc-UdeM
## 13 Levels: 2018-BeauchampRioux-MSc-UdeM ... SWA-Warren
```

```
#which project are the NOT deleted ones missing from? 8 projects
unique(Johanna_missing$project[-which(Johanna_missing$status == "deleted")])
```

```
## [1] CABO-test                    2018-Boucherville
## [3] 2019-Pardo-MSc-UdeM          SWA-Warren
## [5] 2018-BeauchampRioux-MSc-UdeM 2018-Hacker-PhD-UBC
## [7] 2019-Blanchard-MSc-UdeM      2019-Boucherville
## 13 Levels: 2018-BeauchampRioux-MSc-UdeM ... SWA-Warren
```

```
#Verified samples are missing from the CABO-test project
unique(Johanna_missing$project[which(Johanna_missing$status == "verified")])
```

```
## [1] CABO-test
## 13 Levels: 2018-BeauchampRioux-MSc-UdeM ... SWA-Warren
```

```
#Pending samples are missing from CABO-test and 2019-Pardo-MSc-UdeM
unique(Johanna_missing$project[which(Johanna_missing$status == "pending")])
```

```
## [1] CABO-test           2019-Pardo-MSc-UdeM
## 13 Levels: 2018-BeauchampRioux-MSc-UdeM ... SWA-Warren
```

```
#Submitted samples are missing from 7 projects
unique(Johanna_missing$project[which(Johanna_missing$status == "submitted")])
```

```
## [1] 2018-Boucherville            SWA-Warren
## [3] 2018-BeauchampRioux-MSc-UdeM 2018-Hacker-PhD-UBC
## [5] 2019-Blanchard-MSc-UdeM      2019-Pardo-MSc-UdeM
## [7] 2019-Boucherville
## 13 Levels: 2018-BeauchampRioux-MSc-UdeM ... SWA-Warren
```

```
#Johanna's data is missing all 2019-Pardo while Etienne's has 71 samples from 2019-Pardo
"2019-Pardo-MSc-UdeM" %in% unique(Johanna_data$project)
```

```
## [1] FALSE
```

```
"2019-Pardo-MSc-UdeM" %in% unique(Etienne_data$project)
```

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

```
length(Etienne_data$sample_id[which(Etienne_data$project == "2019-Pardo-MSc-UdeM")])
```

```
## [1] 71
```

```r
#Are all the missing samples from Pardo? No, but samples of 71 species are (71 of the 461 are)
length(Johanna_missing$sample_id[which(Johanna_missing$project == "2019-Pardo-MSc-UdeM")])
```

```
## [1] 71
```

```r
length(Johanna_missing$sample_id[-which(Johanna_missing$project == "2019-Pardo-MSc-UdeM")])
```

```
## [1] 390
```

```r
nrow(Johanna_missing)
```

```
## [1] 461
```

Now to compare Etienne's data with the spectral data

```r
# Etienne_data
#
# Shan_spectra
#
# metadata_Shan
#
# species_Shan
#
# Johanna_data

#which samples are missing from full metadata
Etienne_missing_from_Shan <- metadata_Shan[-which(metadata_Shan$sample_id %in% Etienne_data$sample_id),]
Shan_missing_from_Etienne <- Etienne_data[-which(Etienne_data$sample_id %in% metadata_Shan$sample_id),]
Johanna_missing_from_Shan <- metadata_Shan[-which(metadata_Shan$sample_id %in% Johanna_data$sample_id),]

#200 samples missing from Etienne metadata in Shan spectral data
nrow(Etienne_missing_from_Shan)
```

```
## [1] 200
```

```r
#626 samples missing from Johanna metadata in Shan spectral data
nrow(Johanna_missing_from_Shan)
```

```
## [1] 626
```

```r
#which projects are missing - 2017-Dessain-MSc missing from Etienne data
unique(Etienne_missing_from_Shan$project)
```

```
## [1] "2017-Dessain-MSc"
```

```r
unique(Shan_missing_from_Etienne$project)
```

```
## [1] CABO-test                     2018-Girard-MSc-UdeM
## [3] 2018-BeauchampRioux-MSc-UdeM 2019-Crofts-PhD-UdeS
## [5] 2018-Boucherville             2019-Pardo-MSc-UdeM
## [7] 2019-Boucherville
## 13 Levels: 2018-BeauchampRioux-MSc-UdeM ... SWA-Warren
```

```
unique(Johanna_missing_from_Shan$project) # missing from 8 projects
```

```
## [1] "2018-BeauchampRioux-MSc-UdeM" "2019-Blanchard-MSc-UdeM"
## [3] "2018-Boucherville"            "2019-Boucherville"
## [5] "2017-Dessain-MSc"             "2018-Hacker-PhD-UBC"
## [7] "2019-Pardo-MSc-UdeM"          "SWA-Warren"
```

```
#have species IDs for samples missing metadata but not other metadata
is.na(Etienne_missing_from_Shan$species)
```

```
##   [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##  [13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##  [25] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##  [37] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##  [49] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##  [61] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##  [73] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##  [85] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##  [97] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [109] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [121] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [133] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [145] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [157] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [169] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [181] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [193] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
```

```
Johanna_missing_from_Shan$project[which(is.na(Johanna_missing_from_Shan$species))] #included somes samp
```

```
##   [1] "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM"
##   [4] "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM"
##   [7] "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM"
##  [10] "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM"
##  [13] "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM"
##  [16] "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM"
##  [19] "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM"
##  [22] "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM"
##  [25] "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM"
##  [28] "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM"
##  [31] "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM"
##  [34] "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM"
##  [37] "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM"
##  [40] "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM"
##  [43] "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM"
##  [46] "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM"
```

```
##  [49] "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM"
##  [52] "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM" "2019-Pardo-MSc-UdeM"
##  [55] "2019-Pardo-MSc-UdeM" "SWA-Warren"          "SWA-Warren"
##  [58] "SWA-Warren"          "SWA-Warren"          "SWA-Warren"
##  [61] "SWA-Warren"          "SWA-Warren"          "SWA-Warren"
##  [64] "SWA-Warren"          "SWA-Warren"          "SWA-Warren"
##  [67] "SWA-Warren"          "SWA-Warren"          "SWA-Warren"
##  [70] "SWA-Warren"          "SWA-Warren"          "SWA-Warren"
##  [73] "SWA-Warren"          "SWA-Warren"          "SWA-Warren"
##  [76] "SWA-Warren"          "SWA-Warren"          "SWA-Warren"
##  [79] "SWA-Warren"          "SWA-Warren"          "SWA-Warren"
##  [82] "SWA-Warren"          "SWA-Warren"          "SWA-Warren"
##  [85] "SWA-Warren"          "SWA-Warren"          "SWA-Warren"
##  [88] "SWA-Warren"          "SWA-Warren"          "SWA-Warren"
##  [91] "SWA-Warren"          "SWA-Warren"          "SWA-Warren"
##  [94] "SWA-Warren"          "SWA-Warren"          "SWA-Warren"
##  [97] "SWA-Warren"          "SWA-Warren"          "SWA-Warren"
## [100] "SWA-Warren"          "SWA-Warren"          "SWA-Warren"
## [103] "SWA-Warren"          "SWA-Warren"          "SWA-Warren"
## [106] "SWA-Warren"          "SWA-Warren"          "SWA-Warren"
## [109] "SWA-Warren"          "SWA-Warren"
```

It seems that the main issues (now that I have Etienne's data with more of the discrepancies resolved), is that the metadata is missing for the 2017-Dessain-MSc project. For the data availble for me (Johanna) to download from the Fulcrum website, there was more information missing compared to the spectral data, including for samples from 8 projects (with all of the 2019-Pardo project missing), and including samples without species IDs in Shan's spectral data.