

Annexe : Description des différents onglets à renseigner dans le fichier de soumission

La langue recommandée pour saisir les informations est l'Anglais. Le Français est néanmoins accepté.

Gestion des données multiannuelles : Les expérimentations d'un dispositif pérenne sur plusieurs années ou d'un réseau d'expérimentation sur plante annuelles se fait grâce au champs PhenotypingCampaign. Il peut se faire de deux manière. La première avec un fichier par année ; dans ce cas la PhenotypingCampaign doit être renseigné dans l'onglet Trial et laissé vide dans l'onglet Data. La seconde avec un unique fichier pour toutes les années ; dans ce cas la PhenotypingCampaign doit être vide dans l'onglet Trial et renseigné dans l'onglet data.

Les champs en caractères gras sont obligatoires.

1) Onglet Trial et description de l'essai

a) TrialName

- i) Description : Nom de l'essai, texte libre, max 255 caractères. **Caractères alphanumériques, tirets hauts et bas uniquement (pas d'accents, d'espaces, de caractères spéciaux)**. Idéalement une référence à la publication du jeu de données en cours de préparation pour soumissions.
- ii) Exemples: Doligez_et_al_2013, Coupel-Ledru_et_al_2014, ...

b) TrialCode

- i) Description : Identifiant de l'essai. Il doit être unique. Il permet d'ajouter des données à un essai existant : par exemple un premier fichier pour les données 2013 et un deuxième pour les données 2014 utilisant tous les deux le même trialCode. **Caractères alphanumériques, tirets hauts et bas uniquement (pas d'accents, d'espaces, de caractères spéciaux)**. Longueur maximum 255 caractères.
- ii) Exemple : SxG-DoligezEtAl

c) BeginingDate et EndingDate

- i) Description: Date de début de de fin de l'essai. Peuvent être vides. Recommandé pour assurer le lien direct avec Climatik.
- ii) Format: trois possibilités :
 - (1) dd/mm/yyyy
 - (2) mm/yyyy
 - (3) yyyy

d) TrialGoal

- i) Description: Objectifs de l'expérimentation. Texte libre si possible en anglais.
- ii) Exemples : "Study of the genetic determinism of berry weight and seed traits in a grapevine progeny", "Study of the genetic determinism of water use efficiency in a grapevine progeny".

e) LevelStructure

- i) Description : nom de la hiérarchie de niveaux utilisés dans l'onglet data.
- ii) Doit contenir une valeur parmi :
 - (1) Nomenclature ADONIS
 - (a) Vide : pour des données enregistrées sans répétitions
 - (b) BLOC>ParcelleUnitaire>Individu
 - (c) BLOC>SUBBLOC>ParcelleUnitaire>Individu
 - (d) BLOC>ParcelleUnitaire

- (e) BLOC>SUBBLOC>ParcelleUnitaire
- (2) Nomenclatures autre
 - (a) REPLICATION>BLOCK>INDIVIDUAL
 - (b) BLOCK>SUB-BLOCK>PLOT>PLANT
 - (c) REPLICATE>BLOCK>PLOT>REPETITION
- f) PhenotypingCampain
 - i) Description: Nom / étiquette d'un groupe de données au sein de l'essai.
 - ii) Attention : si cette colonne est remplie, il faut que la colonne PhenotypingCampaign de l'onglet Data soit vide.
 - iii) Exemples : « 2012 » pour toutes les observations 2012 ou moyennes 1996-2012 Ou texte libre comme par exemple "Données pluriannuelles".
- g) PMID
 - i) Liste séparée par virgule d'identifiants de publications. Pour l'instant seuls les identifiants PMID (PubMed) sont acceptés. **Si un autre identifiant (DOI, ProDINRA, ...) est nécessaire, contacter Cyril Pommier.**
 - ii) Exemple : 24350702
- h) X et Y
 - i) Description: Nom correspondant aux colonnes X et Y de l'onglet Data.
 - ii) Exemple: X = rang et Y = souche.
- i) PreviousCrop
 - i) Description: plantation existante avant l'essai (permet de décrire une rotation)
 - ii) Exemple: Vigne; friche; luzerne
- j) TrialDesign
 - i) Description: Schéma utilisé pour agencer l'essai
 - ii) Exemple: blocs complets randomisés / Randomized complete blocks
- k) TrialHistory
 - i) Description: Identification des anciens gestionnaires, scientifiques et/ou propriétaires de l'essai
 - ii) Exemple: INRA Colmar, C. Schneider et INRA Bordeaux, D. Forget
- l) ProjectAcronym
 - i) Nom du projet auquel doivent être rattaché les données. Pour contrôler le vocabulaire autant que nécessaire, les projets déjà listés dans la base de donnée GnpIS-Ephesis sont listés dans un **onglet nommé « Static Data » du fichier de soumission**. Si vous avez un doute sur la nécessité de créer un nouveau projet, contactez l'URGI.
- m) Trial Set
 - i) Permet de regrouper plusieurs essais au sein d'un même jeu de données ou DataSet. Le trial set peut être utilisé pour regrouper toutes les expériences de phénotypages réalisées sur une même population en ségrégation, un même panel mais également dans une logique différentes, les expériences réalisées par un réseau d'expérimentateurs sur plusieurs descendances.
 - ii) Exemples : SxG_progeny, RESDUR Network, ...
- n) TrialSet Type
 - i) Description: permet de préciser le type de TrialSet/DataSet.
 - ii) Valeurs possibles : Network, Biological Resources Center (BRC)
- 2) **Onglet Site et description du site sur lequel est positionné l'essai**
 - a) SiteName
 - i) Description: Identification du site expérimental
 - ii) Exemple : La Grande Ferrade
 - b) Latitude
 - i) Description: latitude du site expérimental identifié précédemment

- ii) Exemple: 44.788456
- c) Longitude
 - i) Description: longitude du site expérimental identifié précédemment
 - ii) Exemple: -0.577405
- d) Elevation
 - i) Description: altitude moyenne du site en m
 - ii) Exemple: 10
- e) Exposure
 - i) Description: Exposition dominante du site
 - ii) Exemple: plat, presque plat
- f) Slope
 - i) Description: Pente dominante du dispositif
 - ii) Exemple: null
- g) **CityOrRegionOrCountry**
 - i) Description: Plus proche Ville, région ou pays (mot)
 - ii) Exemple: VILLENAVE_D_ORNON. Ce nom sera vérifié et modifié si nécessaire.
- h) SiteStatus
 - i) Description: statut propriétaire de l'essai. C'est un vocabulaire contrôlé défini par les forestiers pour leurs besoins (voir les trois possibilités ci-dessous).
 - ii) Format : P (= propriété privée), I (= INRA) ou S (= régime spécifique)
 - iii) Exemple: P
 - i) Landowner
 - i) Description: propriétaire du terrain sur lequel est placé l'essai. Important pour les forestiers.
 - ii) Exemples: INRA, IFV,
- j) Comment
 - i) Description: Commentaires sur le site
 - ii) Exemple: parcelle très hétérogène
- k) **MeteoStationCode et MeteoStationNetwork**
 - i) Description: Identifiant de la station météo la plus proche et du réseau auquel elle appartient éventuellement. **Recommandé** pour l'interopérabilité avec Climatik.
 - ii) Exemple MeteoStationCode : 34150002
 - iii) Exemple MeteoStationNetwork : Agroclim-Climatik

2) Onglet Contact et description des contacts en relation avec l'essai

Autant de contacts que nécessaire, une ligne par contact.

- a) Type
 - i) Type du contact (texte libre)
 - ii) Exemples :
 - (1) DataManager
 - (2) DataCollector
 - (3) ContactPerson
 - (4) Scientific coordinator
 - (5) ProjectLeader
- b) Institution : une valeur à prendre dans l'onglet Static Data, colonne InstitutionName. Contactez-nous si nécessaire.

3) Onglet Variables

- a) **OntologyName** : dans le cas de la vigne, il s'agit de « Vitis Trait Ontology », gérée par Eric Duchêne (cf. §3).

- b) **VariableID** : il doit correspondre à l'identifiant de la « Vitis Trait Ontology » (cf. §3).
- c) **ShortName** : **abréviation qui sera ensuite utilisé comme en-tête de colonne dans l'onglet « data »** (voir ci-dessous). Ce label ne devra pas excéder 256 caractères alphanumériques. L'underscore, le point et le tiret sont autorisés mais pas les accents, les espaces, ou les caractères spéciaux.

4) Onglet Plant Material et description des accessions utilisées dans l'essai

L'identification non ambiguë d'une accession repose sur trois champs : HoldingInstitution, Genre et AccessionNumber.

a) **TaxonomicIdentifier**

- i) Nom de la famille. La première lettre est en majuscule
- ii) Ici : Vitaceae

b) **AccessionNumber**

- i) Description: Identifiant unique de l'accession tel qu'enregistré dans la base de données Bioweb-Vigne (cf. Partie 2)
- ii) Exemple: 150Mtp11

c) **AccessionName**

- i) Description: Nom de l'accession présent ici uniquement à titre informatif. Il ne sera pas utilisé pour lier aux accessions déjà insérées de Bioweb-Vigne dans GnpIS, les données en cours de soumission.
- ii) Exemples: Syrah; Grosse Syrah; Syrah Motton

d) **Holding Institution**

- i) Description: Institut dépositaire du matériel végétal. **A reprendre dans l'onglet « static data » du fichier de soumission pour éviter les faux doublons de données.**
- ii) Exemple : Unité Expérimentale du Domaine de Vassal, INRA-Montpellier

e) **LotName**

- i) Description: Identification précise du lot de vigne observé. **Ce code sera celui qui est référencé dans la colonne « LotName » de l'onglet « Data ». Ce numéro de lot est indispensable.**
- ii) Exemple: accession_number-numéro de souche

f) **CollectionsName:**

- i) Description: Nom d'une collection d'accessions. Moyen de définir un sous-ensemble identifié d'accessions.
- ii) Exemples: Grapevine_core_273

g) **CollectionType**

- i) Description: Type de collection. Fait l'objet d'un vocabulaire contrôlé listé dans l'onglet "static data" du fichier de soumission. Contacter S. Durand et T. Lacombe en cas de nécessité de créer une nouvelle catégorie.
- ii) Exemples: NATIONAL, ACTIVE, BASE, CORE, NETWORK, POPULATION, EXPERIMENTAL...

h) **PanelName**

- i) Description: Nom d'une collection de lots. Utilisé en relation avec les modules "génotypage" et "génétique d'association/GWAS" de GnpIS.
- ii) Exemple: Core_273

i) **PanelSize**

- i) Description: Nombre de lots dans le panel
- ii) Exemple: 273

5) Onglet Treatment ou description des modalités de traitement de l'essai

Liste les facteurs et les modalités de traitements (intrants, fongicides, conduite, irrigation, ...) utilisés dans l'expérience. Chaque expérience peut avoir autant de facteurs que nécessaire. La variété est considérée comme un facteur par défaut et seuls les autres facteurs sont précisés dans le fichier.

Exemples de facteurs : traitement, irrigation.

Exemples de modalité pour le facteur traitement : non traité, faible intrant, fort intrant, etc.

Ecrire une modalité par ligne.

Ces champs ne gèrent pas le multilingue. Il vaut mieux favoriser l'anglais.

a) Factor

i) Description: Nom du facteur. **Utilisé comme en-tête dans l'onglet data entre la colonne LotNumber et les colonnes de Level**

ii) Exemples : itk (pour itinéraire technique), water (pour alimentation hydrique)

b) ModeCode

i) Description: Identifiant unique de la modalité du facteur. Il s'agit des valeurs autorisées pour un facteur et qu'on retrouve dans la colonne de ce facteur.

ii) Exemples : pour le facteur "itk", modalités "fi" (faibles intrants phytosanitaires / low pesticide inputs) et "t" (traité/treated)

c) ModeName

i) Description: Nom complet et explicite de la modalité. Affiché sur l'interface web: **favoriser l'anglais.**

ii) Exemples : low pesticide inputs, treated.

d) ModeDescription

i) Description: Champs libre. Peut contenir le protocole.

6) Onglet Data et saisie des résultats de l'essai

Colonne en rouge : à remplir.

Colonne bleue : à ne pas supprimer, mais peuvent être vide.

Les facteurs et les niveaux doivent référencer les Codes déclarés dans les onglets de metadonnées (=tous les onglets précédents).

a) LotName

Doit être un des LotName listés dans l'onglet PlantMaterial.

b) Factors

Le nombre de colonnes factors va de 0 à n. Chaque colonne factor doit correspondre à une des valeurs listées dans la colonne Factors de l'onglet Treatments.

Sur chaque ligne, la valeur correspond à l'un des ModeCode du facteur, déclarés dans l'onglet 'Treatment'.

Exemple: en en tête "Itinéraire Technique" et en ligne, en face des données correspondantes soit "fi", soit "t"

c) Levels

Le nombre de colonnes levels va de 0 à n. Il doit y avoir autant de colonnes levels que de niveaux déclarés dans la colonne LevelStructure de l'onglet Trial. Chaque colonne levels doit correspondre à une des valeurs listées, par exemple pour la LevelStructure REPLICATION>BLOCK>INDIVIDUAL il faut les trois colonnes suivantes : REPLICATION, BLOCK, INDIVIDUAL. Sur chaque ligne il faut entrer la valeur de l'identifiant de l'objet (microparcelle, plante).

Exemple: en tête de colonne: "Bloc" et en ligne 1 pour les données obtenues sur le bloc 1, 2 pour le bloc deux etc...

d) X, Y, X(m), Y(m)

Permet de positionner l'objet qu'on observe (Plante, Souche, MicroParcelle, etc.) dans le dispositif expérimental. Le référentiel à utiliser pour lire ces positions doit être défini par les colonnes X et Y de l'onglet Trial.

Exemples :

XName = X, Yname = Y.

XName = rang, Yname = souche.

e) PhenotypingCampaign

Nom de la campagne de mesure. Peut-être une année ou un ensemble d'années dans le cas de données moyennées.

Exemples :

2001

2001-2002

1956-2012

f) Variables

Il s'agit des variables d'observations listées dans l'onglet Variables.

Le nombre de colonnes variables va de 0 à n. Il doit y avoir autant de colonnes variables que de lignes dans l'onglet Variables. Chaque colonne variable doit correspondre à une des valeurs listées dans la colonne VariableShortName de l'onglet Variable.

Il peut y avoir plusieurs manières de rentrer les données selon qu'on veut garder la date de mesure ou non.

Deux formats de tableaux sont possibles pour cette feuille :

LotName	<factors>	<levels>	X	Y	X(m)	Y(m)	PhenotypingCampain	<variables>	<variables_date>
								Valeur de la variable	Date d'observation de la variable

LotName	<factors>	<levels>	X	Y	X(m)	Y(m)	PhenotypingCampain	date	GDD	<variables>
								Date d'observation des variables suivantes	Growing degree day (degrés jours) des observations des variables suivantes	Valeur de la variable

i) Cas 1 : pas de date d'observation

Exemple de remplissage de la feuille :

LotName	Irrigation	plante	X	Y	X(m)	Y(m)	PhenotypingCampain	BUD	SUGAR_CONTENT
32Mtp0	faible	3					2009	20/06/2009	200

ii) Cas 2 : série temporelle

Pour chaque variable observée, il y a au moins deux dates de mesures. Cette date peut être stockée dans les colonnes Date et GDD (Growing Degree Day : somme degrés jours). La valeur de Date/GDD s'applique à toutes les valeurs renseignées sur la même ligne.

Exemple de remplissage de la feuille :

LotName	Irrigation	plante	X	Y	X(m)	Y(m)	PhenotypingCampain	date	GDD	height	FertCours	Circ
32Mtp0	faible	3					2009				5	

32Mtp0	faible	3					2009	20/05/2009		8		2
32Mtp0	faible	3					2009	05/08/2009		15		3

FertCours= 5, mais sans date.

Height = 8 le 20/05/2009 et 15 le 05/08/2009.

Circ = 2 le 20/05/2009 et 3 le 05/08/2009.

iii) Cas 3 : observation avec traçabilité de la date en colonne

Pour chaque valeur, la date d'observation est renseignée dans la colonne _date correspondante, qui doit immédiatement suivre. Par exemple, la date d'observation de Height est notée dans la colonne Height_date.

Exemple de remplissage de la feuille :

LotName	Irrigation	plante	X	Y	X(m)	Y(m)	PhenotypingCampain	Height	Height_date	Floraison
32Mtp0	faible	3					2009	8	20/05/2009	20/6/2010
32Mtp0	faible	3					2009	15	05/08/2009	

iv) Cas 4 : les observations sont des dates

Cas de la Phenologie. Exemple de remplissage de la feuille.

LotName	Irrigation	plante	X	Y	X(m)	Y(m)	PhenotypingCampain	Floraison	Débourrement
32Mtp0	faible	3					2009	20/6/2010	20/4/2010

1) Onglet Files pour la soumission de documents liés à l'essai

- a) FileName
 - i) Description: nom du fichier
 - ii) Exemples: plan_essai_RESDUR_BDX_2015_675.csv;
itk_details_RESDUR_BDX_2015_675.pdf
- b) Link
 - i) Description: URL où on peut trouver le fichier s'il est déjà stocké de façon stable.
 - ii) Exemple: <https://urgi.versailles.inra.fr/TOTO>
- c) FileType
 - i) Description: format informatique du fichier
 - ii) Exemples: doc, xls, csv, pdf
- d) FileSize
 - i) Description: taille du fichier (nombre attendu dans le champ)
 - ii) Exemple: 36
- e) CreationDate
 - i) Description: date de création du fichier (dd/mm/yyyy)
 - ii) Exemple: 21/05/2015
- f) Description
 - i) Description: champ texte libre permettant de décrire le contenu du fichier.
 - ii) Exemple: plan détaillé de l'essai / Experiment detailed map

2) Onglet ITK

- a) ObjectName
 - i) Description: Objet du dispositif sur lequel a été effectué l'intervention.
 - ii) Exemples: Trial, Plant :37Mtp0, Plot :5
- b) InterventionName
 - i) Description: Intervention réalisée

- ii) Exemples: Taille, désouchage, épendage, traitement fongicide
- c) PhenotypingCampaign
 - i) Description: nom de la campagne d'observation
 - ii) Exemple : 2008
- d) InterventionDate
 - i) Date à laquelle a été réalisée l'intervention.
 - ii) Exemples: 01/01/2008
- e) Comments
- f) Notators
 - i) Description: Personnes ayant réalisé l'intervention.

Source d'information: <https://urgi.versailles.inra.fr/Data/Phenotype/Data-submission>

Référent: Cyril POMMIER