

## I- Informations sur la structure

### *a) Nom de la structure*

CRB FAN de l'UBOCC : Fungi Agri-food eNvironment

### *b) Type de structure*

CRB : Centre de Ressource Biologique

### *c) Identifiants de la structure*

Le CRB FAN fait partie de l'UBOCC qui est un service de l'Université de Brest. Le CRB conserve et distribue plus de 3700 souches de levures et moisissures et de bactéries, issues principalement des activités de recherche du LUBEM (EA-3882 USC 1504 INRAE). Le CRB est hébergé au sein des locaux de l'ESIAB et bénéficie du soutien technique ainsi que de l'expertise du LUBEM (laboratoire support).

### *d) Responsabilités de la structure*

Responsable de l'UBOCC : Amélie Weill

Responsable du CRB FAN : Amélie Weill

Responsable Qualité : Amélie Weill

Data-Manager UBOCC : Amélie Weill avec le soutien de la DSIUN de l'UBO

### *e) Etablissement tutelle*

UBO : Université de Bretagne Occidentale

### *f) Rattachement*

Le CRB FAN est rattaché à l'UBOCC est rattaché à la Vice-Présidence Recherche et Innovation en charge de la commission recherche.

### *g) Financeurs (permettant l'acquisition des jeux de données – hors projet)*

UBO

## II- Informations sur le PGD

Version (cf cartouche)

Rédigé par Amélie Weill

## III- Présentation générale des données

Les données sont associées à des souches de levures, de moisissures, de bactéries ou du matériel biologique (ADN). Elles sont toutes reliées à un n°UBOCC unique. La production des données, la saisie des données, le suivi de la qualité des données ainsi que leur stockage et leur partage sont assurés par l'ensemble des personnels du CRB.

La langue utilisée pour les données accessibles au public via notre catalogue en ligne est l'anglais pour une meilleure compréhension par nos clients internationaux et une plus grande réutilisation de ces données.

*a) Les données du CRB*

Les données produites par le CRB sont stockées sur le serveur CERATOCYSTIS et copiées automatiquement sur l'ordinateur de la responsable du CRB FAN. Les données liées aux souches sont stockées dans le logiciel BioloMICS (cf. VI a)BioloMICS)

*b) Les données recueillies ou produites*

Elles sont fournies au CRB au format demandé, stockées sur le serveur CERATOCYSTIS et copiées automatiquement sur l'ordinateur de la responsable de l'UBOCC. Les autres données sont stockées sur les serveurs CERATOCYSTIS, dans les cahiers de laboratoire ou sur les ordinateurs des personnels des laboratoires. Elles sont également stockées dans BioloMICS.

#### IV- Origine, Mode d'obtention et Droits de propriété intellectuelle des données

*a) Données d'origine de la souche*

- Informations sur la ressource biologique : date d'isolement, personne ayant isolé ressource et laboratoire d'appartenance, méthode d'isolement, collection ouverte ou fermée, visible au catalogue public, condition de diffusion, correspondant scientifique, propriété de la souche, monospore, souche type, identification taxonomique, BSL, type de souche, données culturelles (temps de croissance, pH usuel, pression usuelle, température usuelle, milieux de culture, statut respiratoire)
- Information sur le substrat d'origine : code d'échantillon, code laboratoire, code autre collection, substrat d'origine, pays d'origine du substrat, personne ayant collecté l'échantillon et le laboratoire d'appartenance, date de collecte, Projet associé, données spécifiques aux ressources marines (type de sédiment et profondeur, nom de la mission, station, salinité, profondeur de la colonne d'eau)
- Information sur la caractérisation des ressources : méthode d'identification et validateur, données de séquençage (date, scientifique, nom du projet, nom du gène, nom des primers et leur séquence, séquence, nombre de paires de base, n° d'accession, base de données de référence), analyse des voies métaboliques, séquençage du génome (oui/non), nom du scientifique et son institut, n° d'accession, outil utilisé pour annoter, taille estimée).

Toutes ces données d'origine de la souche sont fournies par les déposants. Le CRB les recueillent dans les fichiers « Fiches souches à déposer ». Ces fichiers sont conservés sur le serveur CERATOCYSTIS de l'UBO ainsi que dans BioloMICS comme décrit dans le paragraphe III a).

Certaines données sont accessibles sur le site internet de l'UBOCC (Cf annexe). Leur accès est régulé comme explique dans le paragraphe V b)

L'ensemble des données d'origine issues des travaux de recherche des laboratoires d'appui est stocké dans :

- des cahiers de laboratoire
- les serveurs de l'UBO (CERATOCYSTIS)
- La base de données des campagnes océanographiques de l'Ifremer.

Toutes les informations liées à l'origine de la souche sont formalisées par une documentation identifiée de manière unique et répertoriée sur la liste des documents par une référence propre. Ces documents permettent de conventionner l'appellation des différentes données et la disponibilité de ces données aux différents acteurs.

*b) Données de caractérisation*

N° UBOCC, méthode d'identification, BSL, Full name et current name

Les ressources déposées en collection sont caractérisées avant leur intégration au sein du CRB et les données sont collectées en amont. Dans le cas d'une caractérisation d'une ressource déjà présente en collection, les données sont stockées dans les bases de données du CRB. Le nom du projet ainsi que le nom du porteur du projet sont reliés aux données de caractérisation et enregistrés dans la base de données. L'ensemble des données de caractérisation issues des travaux de recherche des laboratoires (ou EQUASA) sont stockées dans des cahiers de laboratoire, les serveurs de l'UBO (CERATOCYSTIS). L'information sur la méthode d'identification est accessible sur le catalogue en ligne. Les données de séquençages sont conservées au sein du CRB et ne sont diffusables hors UBO qu'après accord du laboratoire ayant réalisé le séquençage. Elles sont accessibles sur le site BioAware du CRB que par les membres des laboratoires (cf annexe). Les données sur les génomes sont conservées par les laboratoires propriétaires, le CRB ne garde que l'information oui/non.

*c) Données de distribution*

Elles sont conservées dans un MTA qui est archivé dans CERATOCYSTIS et dans BioloMICS. Le droit sur ces données appartient au CRB et elles sont confidentielles et non diffusables en dehors du CRB.

*d) Format et saisie des données*

L'ensemble des données saisies dans les bases de données est conservé par BioAware.

**Type et Format de données**

Pour l'acquisition des souches, leur suivi, leur distribution ainsi que pour répondre aux demandes de nos utilisateurs, diverses données sont collectées et/ou générées par les membres du CRB. Ces données sont sous différents formats :

Type de données	Format
Format texte	Archive papier, .doc, .docx, .txt, .odt., .pdf
Format table calcul	.xls, .xlsx, .csv, .txt, .odf
Base de données	.csv
Image	.jpeg, .jpg, .tiff, .png, .bmp, .pdf
Sequence	. ab1*, .pdf*, .fasta*, .fastq, . bam

Les données sont conservées au format papier par certains services centraux de l'UBO et par la gestionnaire administrative de l'UBOCC. Les documents sont conservés sous format pdf au sein du serveur CERATOCYSTIS de l'UBO.

V- Cadre juridique

*a) Les modalités de diffusion de l'information (RGPD)*

Le CRB se base sur la fiche référentielle : Organisation et Gestion de collections de ressources biologiques (matériel vivant ou inerte)

Afin d'informer ses utilisateurs, l'UBOCC a imposé systématiquement sur ses documents (devis, MTA, mails...) la phrase suivante :

*Pour le suivi des dépôts et des distributions des souches, l'UBOCC a besoin de collecter certaines données personnelles (nom, prénom, civilité, adresse électronique, adresse postale, coordonnées téléphoniques) de ses clients et utilisateurs. Conformément au règlement général sur la protection des données personnelles (RGPD), le client/utilisateur bénéficie d'un droit d'accès, de rectification et, lorsque la réglementation ne s'y oppose pas, de suppression des données personnelles qui le concernent. Pour plus d'information, vous pouvez nous contacter à l'adresse [ubocc@univ-brest.fr](mailto:ubocc@univ-brest.fr)*

*In order to monitor deposit and distribution of strains, UBOCC needs to collect certain personal data (surname, first name, civility, e-mail address, postal address, telephone) from its customers and users. In accordance with the general regulations on the protection of personal data (RGPD), the customer/user has the right to access, rectify and, where regulations do not preclude this, delete personal data. For additional information, do not hesitate to contact us at [ubocc@univ-brest.fr](mailto:ubocc@univ-brest.fr)*

### *b) Protocole Nagoya*

La diffusion et l'utilisation ainsi que la valorisation des organismes détenus au sein du CRB respectent la réglementation Européenne en vigueur sur la gestion des ressources génétiques (mise en œuvre du protocole de Nagoya au travers de la réglementation Accès et Partage des Avantages : APA).

Législation du pays fournisseur ([ABSCH](#))

Pour les souches qui proviennent de pays qui n'ont pas ratifiés le protocole de Nagoya, la législation en vigueur du pays est appliquée.

## VI- Stockage, Archivage et Sécurité des données

### *a) BioMICS*

Après avoir développé une base de données avec les informaticiens de l'UBO, l'UBOCC a choisi de migrer sa base de données sur une solution commerciale fournie par BioAware. C'est un système de gestion de données biologiques utilisé par de nombreux acteurs dans les domaines de la biotechnologie, de l'industrie pharmaceutique, de l'alimentation, des services de santé publique, des institutions académiques, Universités, collections et musées. Il permet de développer la base de données ainsi que le catalogue en ligne sans nécessiter le recrutement d'un informaticien.

BioMICS est un logiciel commercial produit par la société BioAware dont le siège est basé en Belgique. C'est un « software as a service » (SaaS) ou logiciel en tant que service, c'est à dire qu'il est installé sur des serveurs distants plutôt que sur la machine de l'utilisateur. L'accès au logiciel en version desktop se fait en utilisant le protocole RDP (Remote Desktop Protocol). La licence (annuelle) est réservée à un compte utilisateur pour une et une seule connexion à la fois. L'interface utilisateur de BioMICS.desktop est « user-friendly » et inspirée de l'environnement Microsoft permettant ainsi à toute personne, informaticienne ou non, de s'appropriier le logiciel. BioMICS enrichi son offre avec BioMICS.web qui permet le déploiement d'une interface web à partir de la base de données. Pour l'UBOCC cette interface est disponible à l'adresse suivante (<https://ubocc.bio-aware.com/page/Home>). Cette fonctionnalité permet de sélectionner et de rendre disponible en ligne les ressources souhaitées au client d'une manière sécurisée et à jour.

Pour les aspects sécurité des données, l'entreprise sous-traite des tests de pénétrations à Intruder (<https://www.intruder.io>), une société experte en audit sécurité des infrastructures informatiques. Tous les serveurs sont protégés par la solution SentinelOne. Les serveurs sont sauvegardés intégralement tous les jours et les copies conservées une semaine. Tout le système est dupliqué ce qui garantit une haute disponibilité. Une version de sauvegarde activable en 2 heures maximum en cas de

désastre sur le site principal (Site Recovery option) est également disponible dans un autre datacenter du groupe Microsoft (North Europe). Une archive de chaque database client est créée quotidiennement et sauvegardée sous format compressée (backup MySql) pendant une période de 30 jours sur les serveurs de BioAware mais est également déposée quotidiennement dans un DropBox et mis à la disposition du client (dans DropBox, les 5 versions les plus récentes sont gardées). Si le partenariat s'arrête, la société s'engage à restituer l'intégralité des données sous format ouvert (json/csv/Excel/MySQL).

Concernant l'hébergement des serveurs et donc des données, BioAware utilise les solutions Microsoft-Azure. La zone géographique de résidence est située en Europe (« West Europe – Netherlands).

Pour les aspects d'administration, le SI BioloMICS enregistrant tous les changements, un historique est consultable et toutes les actions, y compris celles entraînant des modifications de la structure de la base de données peuvent être annulées par les administrateurs de la Base de données. Si le CRB rencontre un problème au niveau de la base, BioAware résout le(s) problème(s). Dans certains cas rares ou très compliqués, la restauration de la base à une date précise est possible en utilisant les sauvegardes disponibles. Il a été convenu de demander une sauvegarde de la BDD BioloMICS tous les 4 mois qui sera stockée sur CERATOCYSTIS. Nous conserverons toujours les deux dernières sauvegardes.

Les données stockées dans notre base de données BioloMICS sont accessibles à tous ceux qui ont les droits. L'obtention des droits est gérée par la responsable du CRB FAN qui en est l'administratrice. Seule cette dernière a accès à l'ensemble des données de la base de données. Cela se fait par un login/mot de passe commun + confirmation DUO sur un téléphone pour accéder au serveur virtuel puis par une connexion à la base de données protégée un login et mot de passe.

### *b) Les données numériques*

Elles sont stockées sur le serveur CERATOCYSTIS avec une sauvegarde totale mensuelle, conservée 12 mois et une sauvegarde incrémentale quotidienne, conservée 2 mois. Une sauvegarde incrémentale quotidienne est réalisée sur l'ordinateur de la responsable de l'UBOCC. Les données stockées sur CERATOCYSTIS sont accessibles aux seuls membres du CRB et au service informatique de l'UBO. Le serveur est sur onduleur pour limiter les coupures suite à une coupure électrique brève.

### *c) Les formats de conservation des documents*

Les données informatiques sont conservées dans la base de données, sur le serveur CERATOCYSTIS ou sur le DRIVE via des identifiants UBO.

### *d) Les ressources biologiques*

Les souches déposées en attente d'entrer en collection sont stockées à température ambiante ou dans un frigo à 4°C dans une pièce dont l'accès est restreint. Les souches entrées en collection sont conservées à -80°C dans des congélateurs stockés dans deux bâtiments : celui de l'ESIAB et celui de l'IUEM. Dans le bâtiment de l'ESIAB, les congélateurs sont stockés dans une pièce fermée à clef, accessible aux détenteurs d'un pass général. Dans le bâtiment de l'IUEM, les congélateurs sont fermés à clefs et accessibles aux personnels de l'UBOCC. Les souches lyophilisées sont conservées dans le bâtiment de l'ESIAB, dans une pièce dont l'accès est restreint aux détenteurs d'un pass général.

*e) Données sur les lieux et les conditions de stockages*

Elles sont stockées dans BioloMICS dans un module de stockage et dans une table dédiée. L'accès à ces données est réservé à la responsable du CRB FAN comme décrit dans le paragraphe « données numériques ».

*f) Données détruites*

Si le propriétaire d'une ressource souhaite que celle-ci ne soit plus référencée dans le CRB alors la ressource sera inscrite comme perdue dans la base, la ressource détruite et les données associées ne seront plus visibles sur le catalogue en ligne. Les données sur les cahiers de laboratoire doivent être conservées minimum 25 ans et les données afférentes à ces personnes notées dans les cahiers de laboratoires seront alors supprimées dans un délai minimal de 25 ans. Les documents officiels (contrat, MTA...) eux ne seront pas détruits. Les codes de stockages sont réassignés à une nouvelle ressource lorsque la souche n'est plus référencée dans la collection.

*g) Coût de stockage et de partage des données*

Les coûts de fonctionnement et de maintenance des serveurs hébergés par l'UBO sont pris en charge par l'UBO. Pour la base de données, le module stockage et le catalogue en ligne géré par la société Bio-Aware, la licence a un coût de 4286,4 € HT annuel.

## VII- Partage et réutilisation des données

Le CRB ne partage pas ses données. Les données passeports (N°UBOCC / Identification taxonomique / date d'isolement / Pays d'origine / Substrat d'origine / Autre collection / méthode de validation de l'identification taxonomique / Niveau de risque biologique / Souche type) Monospore / température et durée d'incubation : génome séquencé (oui/non) sont accessibles à tous sans identification préalable.

Les informations fournies par le LUBEM sont accessibles au LUBEM après identification sur le site web.

Les autres données ne sont accessibles qu'au personnel du CRB FAN