URGI QUALITY MANUAL

Unité de Recherche Génomique-Info
Versailles-Grignon INRA Center

Quality Management System according to ISO 9001 v.2008 standard

URGI is certified since November 2012

Table of content

1 INRA and URGI

1.1 The French National Institute for Agricultural Research INRA

Ranked number one agricultural institute in Europe and number two in the world, the French National Institute for Agricultural Research (abbreviated as INRA) carries out mission-oriented research for high-quality and healthy food, competitive and sustainable agriculture and a preserved and valorized environment.
INRA researches are guided by developments in scientific fields and focus on worldwide challenges related to food and nutrition, the environment and land use facing the world of agriculture and agronomics today. Challenges such as climate change, human nutrition, competition between food and non-food crops, the exhaustion of fossil resources and appropriate land management put agronomists in a position to generate compatible economic, social and environmental development. INRA produces fundamental knowledge that leads to innovation and know-how for society. INRA lends its expertise to public decision-making.

Read more information on http://www.inra.fr/.

1.2 The Versailles-Grignon research center

The Versailles-Grignon research center is one of the 19 INRA centers. It counts 26 research units distributed over 7 geographical sites.

- Evry
- Gif-sur-Yvette
- Ivry-sur-Seine
- Marne-la-Vallée
- Paris: AgroParisTech, Ecole Normale Supérieure, Ecole d'Economie de Paris
- Thiverval-Grignon
- Versailles

About 1400 people are working in this center (about 1/10 of the INRA total). It counts 230 researchers, 380 engineers, 420 technicians and assistants, 150 researchers and lecturers from other research institutes or universities, and 260 doctoral and post doctoral students (40% of foreign students).

The research activities are structured according to three areas:

- Plant genomics and integrative biology.
- Innovation, agricultural production and environment.
- Food quality.

Read more information on http://www.versailles-grignon.inra.fr/.

1.3 The URGI Unit

The "Unité de Recherche en Génomique Info" (URGI, UR1164) is an INRA research unit from the Biology and plant breeding department (BAP). The research activity covers data integration as well as genome structure and dynamics.

1.3.1 The “Genome analysis” team

We develop tools and acquire knowledge on genome structure, evolution and functioning.

Our research is focused on transposable elements and repeats that appear today to play an important role in genome evolution and its functioning. Indeed, they are major actors of genome size and structure. Transposable elements, viewed for a long time as intra-genomic parasites, are also considered, today, to be at the origin of many essential biological functions such as the immune system, the centromeres, the heterochromatin or the telomerases.

Knowledge and data gained are added to the information system increasing the value of deposited data. The methods and tools developed by this research activity are proposed to the platform users.

1.3.2 The “Information system” team

The access to truly FAIR (Findable, Accessible, Interoperable and Reusable) data remains a challenge for plant biologists and hampers the emergence of new research hypothesis or results from the integration of existing data. URGI develops tools and approaches to make both data and information systems interoperable and tests them through its information system, GnpIS, with the aim to contribute to help the plant research community to cope with this challenge.

Read more information on https://urgi.versailles.inra.fr/.

1.4 The URGI bioinformatics platform

The URGI unit hosts a platform belonging to the National Institute of Bioinformatics (IFB), the French node of the Elixir European infrastructure for computational biology. It belongs to the “Ile-de-France” node of IFB, called Aplibio. URGI platform was labeled as a “strategic bioinformatics platform” at INRA level (by the CNOC committee) and at national level by GIS IBISA, an entity that coordinates a national network of platforms in life sciences.

Platform services cover data management, database design, software engineering, software hosting, data integration and training.

The platform develops and maintains a modular and interoperable information system for plant and pest genomics, called GnpIS, composed of databases, databanks and sequence analysis tools. It enables scientists to mine genomic and genetic data, to extract valuable information on genes of agronomical interest and on genome structure and evolution. GnpIS stores data from several species such as *Arabidopsis*, rice, maize, wheat, grape … but also some insects and fungi. It has been selected by several genome consortia to be an international data warehouse (wheat, grape, and several fungi).

The platform also develops a complete annotation system to annotate genes and repeats, to distribute manual curations among dispersed experts and to store and to visualize annotations. This system is a part of GnpIS, sharing all its search capabilities.

Information systems such as GnpIS play today a central role in the analysis of the ever increasing flow of new genomic data. But new high throughput technologies are very challenging for data integration. Storage is not the only difficulty. Fast data access through queries is even more challenging as this is critical for researchers to efficiently work with their data. To face this new challenge, URGI improves the GnpIS architecture.

3 The URGI unit

3.1 Organization

The Unit counts more than 25 people who dedicate around 22 plain time equivalent positions to platform activities. We have 2 thematic teams and 2 functional “agile” teams.

The thematic teams are:

- Information System and Data Integration
- Genome Analysis

The functional “agile” teams follow agile methodologies. They are organized around technical skills to facilitate know-how sharing. They are:

- Pipeline development
- Information system development and Data management

Any staff can belong to one thematic and one or two functional “agile” teams. The time spent in each team depends on his/her missions.

See up-to-date version on https://urgi.versailles.inra.fr/About-us/Team.

3.2 Workspace, equipments

The servers are hosted in the INRA datacenter near Paris. Backups are done with robotic data cartridges. Secure access through SSL (Secure Shell Layer: https, ssh) and firewalls is provided.

All servers are connected to a high band width network. The internet access is provided through a fiber channel link (10Gb/s) to the global network RENATER (National Telecommunication network for Technology Education and Research), http://www.renater.fr.

The platform offers high user accessibility.

URGI servers are dedicated to bioinformatics: web site, high-throughput computing, and data storage:

- **High- performance cluster**
  - 864 cores
  - 1 master node
  - 1 submission node
  - 77 execution nodes
    - 16 nodes 16 GB RAM
    - 56 nodes 48 GB RAM
    - 4 nodes 96 GB RAM
    - 1 node 512 GB RAM

- **High-performance storage**
  - For Unix, database & cluster
  - 200 TB

- **Virtualization**
  - 3 servers
  - 88 CPUs, 200 GB RAM, 6 TB storage

3.3 Platform supervision

As many platforms, a user scientific committee helps the platform in its decisions since 2005. It is composed of scientists who represent different areas of the genomics field. This committee is responsible of the definition of large guidelines and priorities and gives feedback for the platform activities (1 meeting each year).

Platform steering is the responsibility of the head of unit as well as the heads of the 3 departments: biology and plant breeding department (BAP), plant health department (SPE) and forest, grassland and freshwater ecology department (EFPA).

In the same state of mind, to help most projects to be conducted, a user scientific committee and a steering committee are created to guarantee a good fit between developments and biologists needs.

3.4 Clients / Users

Users are world-wide researchers working in plant sciences. We have different types of user according to the level of service provided:

- **Project partners** are those involved in a scientific collaboration with at least a unit member. These collaborations maybe or not, formalized in a project funded by a funding agency (ANR, EC …). In the frame of the project, URGI may design databases, develop tools, or analyze genomic sequences. The duration of the project may exceed 3 man months.

- **Institutional partners** are members of the INRA departments supporting the platform. Those are from BAP, SPE and EFPA departments. They access all the services through a dedicated call for proposal. Accepted projects cannot exceed 3 man months.
- **Account users** have a login and a password to access restricted access data in our databases, run computer intensive tools, access data banks, or use a Unix shell for computation. These users may be institutional partners or project partners.
- **Anonymous users** use our services through the web portal without signing in with a login and a password. They can access public data in our databases through their web interfaces, run some analyses through the web portal, and download tools or data available through our web site (https://urgi.versailles.inra.fr).

Communication is provided by email aliases and mailing lists. In particular we manage the 2 following lists:

- urgi-contact@inra.fr: for external people who want to have a contact with URGI team (the main URGI email),
- urgi-support@inra.fr: for account users who have questions on platform resources

3.5 Platform charter

The URGI platform offers services. Requests to the platform must be related to one of the following services:

- **Software development:** Databases, information system, sequence analysis tools, pipelines.
- **Data management:** Database insertions, data integration/Interoperability, banks, blast servers.
- **Annotations:** Genes and repeats automatic annotations (structure & functions), distributed manual curation.
- **Bioanalysis:** Genome analysis (structure, function, evolution), high-throughput sequence analysis (SNP calling, RNAseq).
- **Expertise:** Software development, genome analysis.

These services can be requested in the frame of:

- A global need from plant breeding (BAP), plant health (SPE), or ecology (EFPA) departments as an institutional or project partner.
- A call for proposal from a funding agency (ANR, EC …) as a project partner.
- URGI service offering. Restricted to INRA plant breeding (BAP), plant health (SPE), and ecology (EFPA) departments as an institutional partner.

Selection criteria are by priority order:

1. **INRA plant breeding (BAP) and plant health (SPE) departments' priorities:** Species of interest, QTL detection, genetic resources, genomic selection, and genome analysis.
2. **URGI unit research topics:** Transposable elements, genome structure, function, genome dynamics and evolution.
3. **Transferable knowledge**
4. **Generic tools**

4 Quality System Management

4.1 Application domain

This quality manual describes URGI activities and arrangements that are implemented to meet the requirements of ISO 9001 standard.
The scope of application concerned and described in this quality manual allows the unit to fulfill its missions as "a bioinformatics research unit for studies of plant and pest genetics and genomics". It concerns the following activities realized at URGI:

- Manage Unit
- Analyse genomes
- Develop an Information System
- Offer Platform services

4.2 Quality policy and responsibilities

The quality policy statement is available in its entirety in the document QM*_1_Quality_policy*, written by Hadi Quesneville, head of URGI unit. Here is a summary:

"Quality objectives
URGI has established a quality control process to meet the requirements and specificities of a research unit hosting a bioinformatics platform. As part of its activities, the management team is committed to ensure competences and availability of personnel and equipments. The staff is committed to satisfy its users and to continuously improve their satisfaction. Each member is engaged in the quality process to achieve the following objectives:

0. To improve the control of Unit processes
1. To improve the quality of the offered services
    a. To provide robust software and databases
    b. To provide robust computer resources
    c. To provide up-to-date consistent data
2. To manage bioinformatics for large collaborative projects
3. To improve user support
4. To improve staff skills
5. To improve the national and international visibility

Head responsibilities
As head of the URGI, I have to integrate the requests from the unit and from the INRA. I take the pledge to:

- promote a culture of quality to improve the working methods,
- nominate a quality moderator to implement the quality policy and assist him/her,
- ensure that human and material resources are sufficient to carry out the unit missions,
- ensure the existence of proper and regular consultations with platform members to ensure the best fit of the platform missions to the demands of users.

Hadi Quesneville, Head of URGI"

4.3 Management principles

The quality management system (QMS) is established to meet the user needs and expectations and continuously improve their level of satisfaction. It is periodically reviewed and updated in order to follow changing needs and expectations.

The PDCA (Plan, Do, Check, Act or Deming wheel of progress) is applied to all processes that

are built according to four major phases: plan, deploy, understand, and improve. The application of its principles allows better management, by measuring and trying to understand the differences in order to progress and to move towards continuous improvement.
In this context the capitalization of these achievements is based on the URGI quality documents.

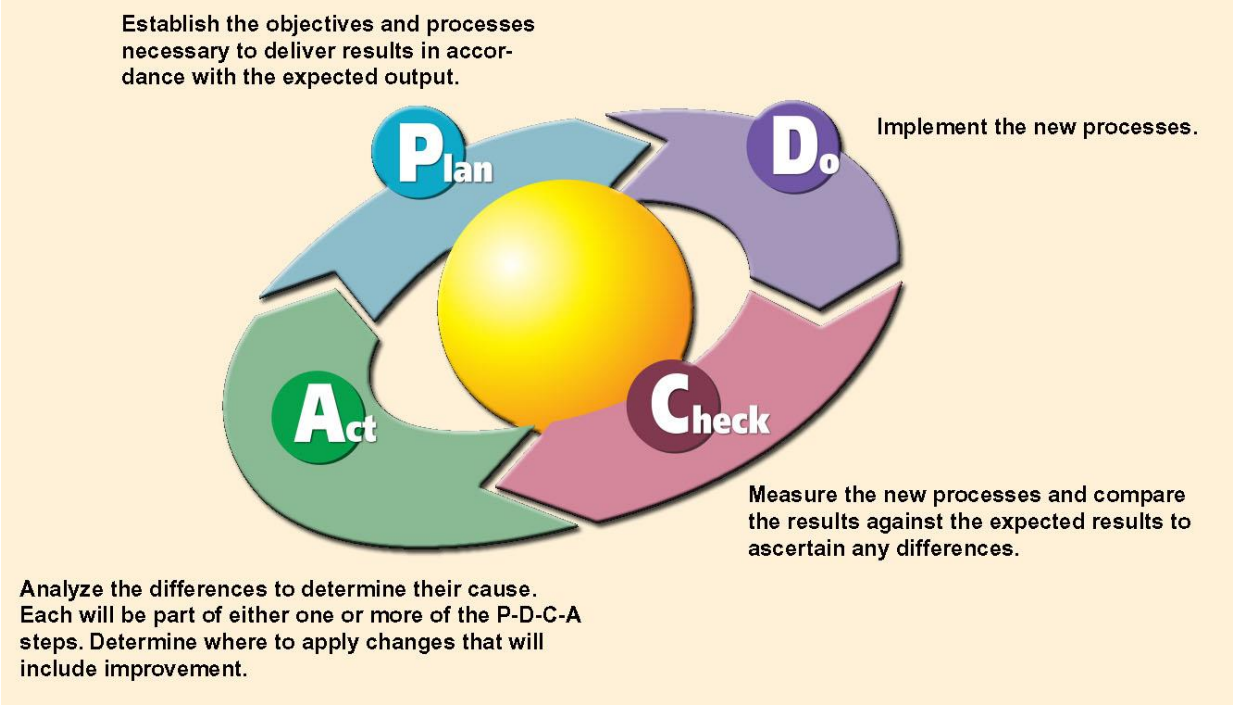
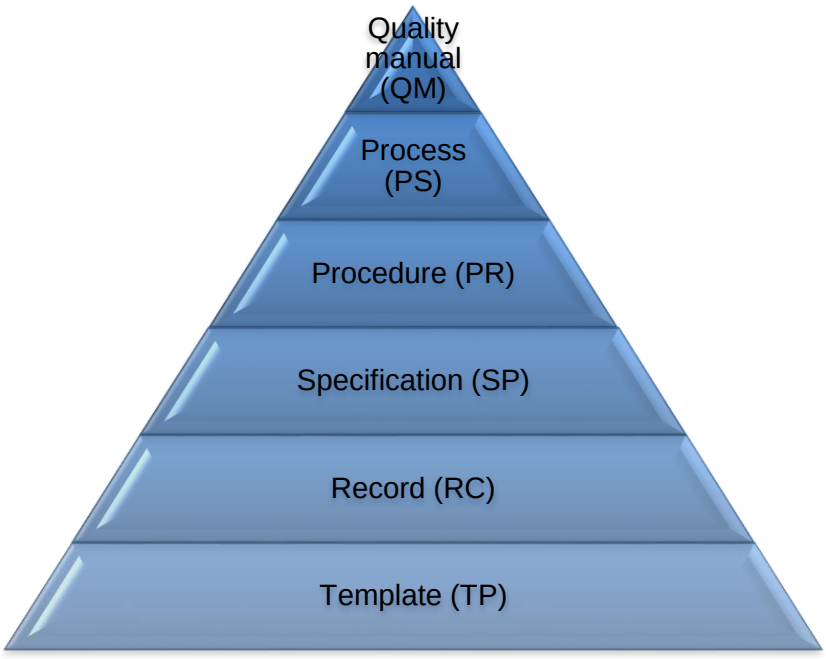


Figure 1: The Deming wheel or PDCA

4.4 Documentation management

The quality management system relies on a documentary system having a pyramidal structure and composed of 6 levels:

- **Quality manual** (QM): Synthetic vision of URGI organization and measures taken to satisfy with the ISO 9001:2008 standard requirements. This document.
- **Process** (PS): Set of correlated and interactive activities which transform input elements into output elements.
- **Procedure** (PR): Description of one or more activities specifying “who does what, when”.
- **Specification** (SP): Chronological description of the tasks of an activity, specifying “how” to do it.
- **Record** (RC): Proof of the realization of an activity, insuring traceability.
- **Template** (TP): Framework giving the structure for a new document.



All documents of the quality system follow rules for their creation, update and diffusion, which are defined in *PR_QAL_001_Documentation_management* procedure. This procedure also defines modes of identification, use, classification and archiving of URGI records.
All applicable documents, records and templates are listed in *RC_QAL_003_Quality_document_list*. This record details the identifier, the name and the domain of each document.

4.5 Processes

Key processes have been identified, modeled and then link together in order to manage them the best way. The process representation in the process map and in the process identity cards make it possible to structure the activities, to identify the quality key points and to implement the continuous improvement on the basis of an appropriate documentation.

4.5.1 Types of process

The quality management system is organized in processes; identified processes have been classified into 3 categories:

- **Two steering processes**:

These processes describe the actions needed for management.

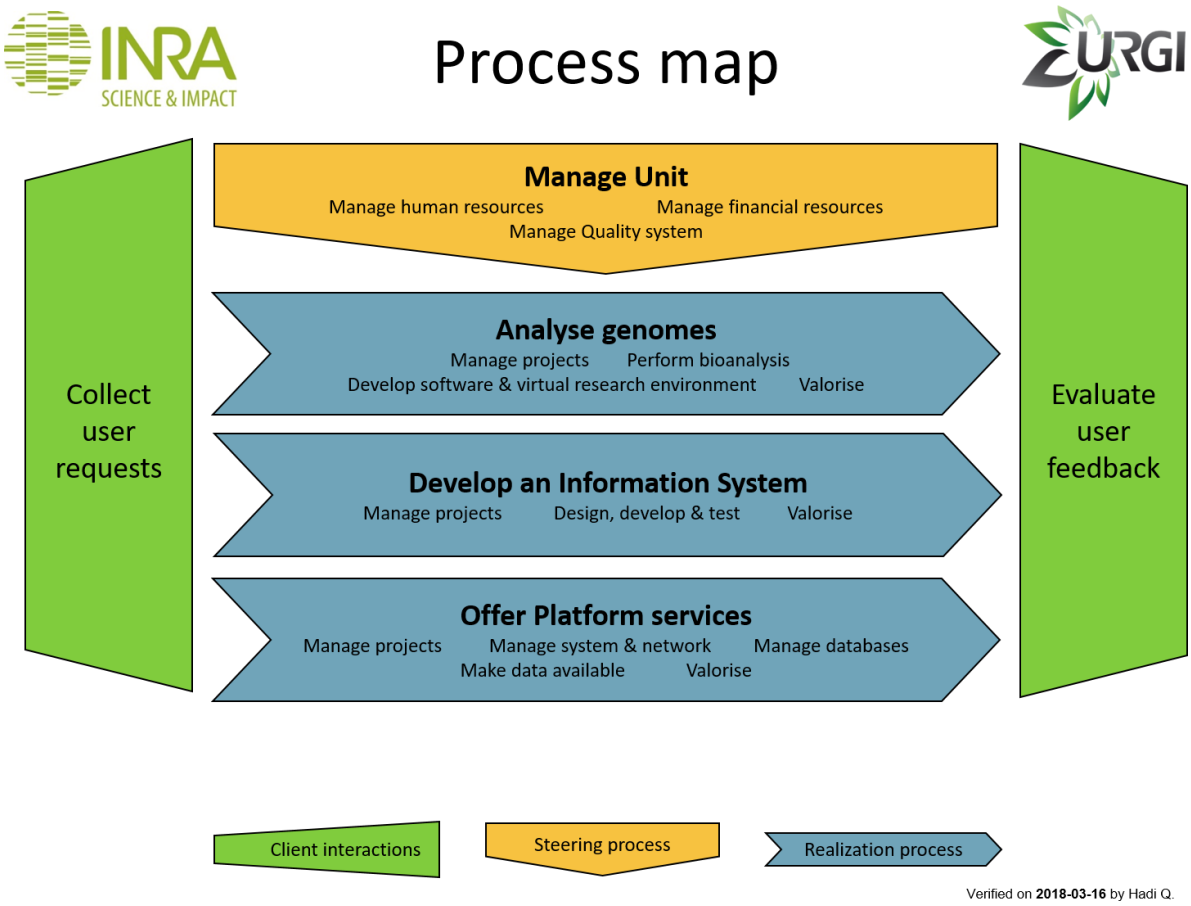
- **Four realization processes**:

These processes describe the business activities of URGI. They are initiated at the request of the client and lead to the realization of a product.

- **Four support processes**:

These processes describe activities that support realization processes. They are necessary for the proper conduct of business processes; their clients are internal.

4.5.2 Process map



4.5.3 Process pilots

Each process is piloted by a process pilot which role is to ensure the **deployment of the objectives** fixed by the Unit direction and the **achievement of results** for its process. He/she has also to ensure the effectiveness of its process, thanks to indicators, user satisfaction. In case of nonconformity or environment evolution, he/she sets up **improvement** actions.

Here is the list of the pilots:

- Manage Unit: Unit Director
- Analyse genomes: Leader of “Genome analysis” team
- Develop an Information System: Leader of “Information System and Data Integration” team
- Offer Platform services: Platform operating manager

4.5.4 Process actors and clients

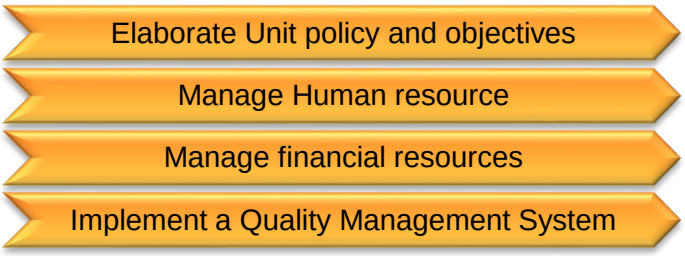
A process actor realizes one or more activities for this process, following the process procedures and specifications. He/she assists to the process review.

A process client asks the process actors to do an activity from this process, and get the result. He/she doesn’t assist to the process review.

4.5.5 Steering process

4.5.5.1Manage Unit (MGT)

The Head of Unit integrates requirements from users, platform User Scientific Committee (CSU), INRA directives and regulatory requirements applicable to URGI in order to define the URGI quality policy. During management review, it can be redefined depending on new requirements.
This process describes the Unit decisional workflow using quality management, as follows:
Ensure efficient management of the Unit, the Quality Management System, Human resources and Financial resources.
The complete process is modeled and described in the document *PS_MGT_Manage_Unit.*



4.5.6 Realization processes

4.5.6.1Analyse genomes (ANAGEN)

In the frame of a project, the bioanalysts perform computational bioanalysis in order to obtain a publication from an original research question. This process aims to improve traceability and reproducibility of bioanalysis. It is modeled as follows:
It is detailed in the document *PS_ANAGEN_Analyse_genomes*.

4.5.6.2Develop an Information System (IS)

In a frame of a project, the “Information System” development team realizes software from client requirements. This process includes all the steps from requirement collect to maintenance, passing by design, development and test of the applications. The complete analysis of this process is available in the document
PS_IS_Information_System.

4.5.6.3Offer Platform services (PF)

This process concerns the management of the URGI platform including its technical maintenance. The complete analysis of this process is available in the document
PS_PF_Platform.

4.6 Measurement, analysis, improvement

The URGI quality system is regularly evaluated and improved through several means: monthly Unit meeting, process reviews, management reviews, internal audits, analysis of nonconformities...

4.6.1 Process and management reviews

Each process is monitored via **indicators** resulted from objectives from quality policy. These indicators are followed by each process pilot and are stored in *RC_QAL_001_Indicators* for each process.

An analysis of each process is done at least once a year with the pilot and all actors to make sure that this process is efficient and to determine improvement opportunities. A summary is done during the **annual management review** where the Head of Unit can validate improvement action plan.

4.6.2 Collect user requests and evaluate user feedback

All Unit members listen daily to their users. Moreover, platform **user committee** and project user committees are the ideal place to listen to users and evaluate their feedback.

On one hand, a **unique email address** centralizes user requests which are treated by the most concerned URGI member. The document *PR_CLI_001_Manage_user_requests* details this procedure.
URGI service offers allow formalizing user demands through small project writing.

On the other hand, **punctual surveys** are organized to evaluate user satisfaction. They are focused on a project, a realization or a training. They are analyzed during management review in order to **improve relationship with users** and **define new orientations**.

4.6.3 Internal audits

The URGI quality management system is audited at least every two years, according to the procedure *PR_QAL_003_Manage_internal_quality_audit*. The objectives are to **check conformity** of the quality system with ISO 9001 standard requirements and to make sure that **continuous improvement is effective.**

4.6.4 Nonconformities and improvement

Detection of a problem (real or potential) launches the **creation and monitoring of nonconformity**, in order to identify, correct the problem and prevent it to re-appear. An analysis of all these nonconformities is done at least during the annual management review. This organization is described in the procedure *PR_QAL_002_Manage_nonconformities,* which details amongst others implementation of corrective and preventive actions.

4.7 Hygiene and security

A URGI member is dedicated to prevention implementation. She has followed dedicated training and is implicated in a working group at INRA Versailles. URGI applies all INRA procedures and a risk analysis has been done.
All events (alarms, sickness …) are recorded in the unit registry (note book and electronic file).

4.8 Tools for quality management

Two main software are used to manage quality documents and traceability. The use of these software is integrated in the daily work of each URGI member, not only for quality, but also for other activities.

- **SharePoint** is an electronic document manager. This web application stores and archives all the documents of the URGI Unit, including quality documents.
- **JIRA** is an electronic task manager. It is Initially used to track bugs, tasks, and entire software development projects, but is also suitable to trace other activities, including user request, nonconformities, …

5 Glossary

ANR: Agence Nationale pour la Recherche
BAP: Département INRA Biologie et Amélioration des Plantes
CSU: User Scientific Committee
EC: European Commission
EFPA: Département INRA Écologie des Forêts, Prairies et milieux Aquatiques
QMS: Quality management system
SPE: Département INRA Santé des Plantes et Environnement