Description of Major Research Tasks

1. Comparative Safety Evaluation of Breeding Approaches and Production Practices Deploying High- and Low- Input Systems

Objectives:
- To identify risks associated with specific breeding approaches.
- To compare risks inherent in high- and low- production systems.
- To explore profiling methods to differentiate risks.
- To develop comparative databases for risk assessment of foods produced in the different systems.
- To evaluate and define the compositional variation of raw materials in the context of a "history of safe use".
- To actively make use of completed and currently on-going research based on these specified in the earlier sections of this technical annex.

Selected crops: Potato (*Solanum tuberosum*) and Maize (*Zea mays*)

Co-ordinator:
Partner 2, SCRI

	Partners	**Contribution**
1	RIKILT (Inst. Food Safety; NL)	Transcriptomics
2	SCRI (Scott. Crop Res. Inst.; UK)	Metabolomics (GCMS; LCMS)
3	TUM (Techn. Univ. Munich; DE)	Metabolomics (GC & GCMS); Proximates
4	UKU (Univ. Kuopio; FIN)	Proteomics and Sub Proteomics
5	IHAR (Plant Breed. Accl. Inst.; PL)	Pathology; Proximates
6	ISS (Inst. Pub. Health; I)	Mycotoxins
7	CSIR (Counc. Sci. Ind. Res.; SA)	Metabolomics (NMR); Pathology
8	BiOSS (Biomath. Stat. Scot. Res.; UK)	Statistics
9	ICGR (Inst. Crop Germpl. Reso.; CHN)	Mycotoxins (see above), proteomics

Deliverables:
- Training and harmonisation workshops (four) for analytical approaches of sample and data handling.
- Ring testing of shared analytical procedures (harmonisation, validation) based on among others criteria as published by the Commission, CEN and ISO.
- Statistical analysis approaches (tested on existing profiling data).
- Identification of potential emerging risks (new matrices) associated with plant pathogens (including mycotoxins).
- Production of appropriate wide range of crop samples for comparative analysis (first set at estimated 3 months; second set at estimated 12 months).
- Establishment of first comparative databases for profiling and proximate analysis.

Workplan

1. Plant Materials for Comparative Studies

The comparative analysis is the first of its kind attempted at the scale envisaged. The partnership within WP1 will provide access to wide ranging population of important and relevant GM and non-GM germplasm and access, through commercial subcontractors, to samples grown under high and low input system. A key driver will be the need to assess the extent of variation (compositional, potential food borne toxins – see above issue on mycotoxins) to which the consumer might be exposed. Assessment of variation will concentrate on 1) Genetic Variation (including that potentially caused by the breeding method (including GM) 2) Variation due to crop production and storage practices.

1.1 Genetic Variation:

Potato

1.1.1 Cultivated potato ("natural" variation). The objective here is to develop an empirical database using profiling approaches to illustrate existing variation in cultivated potato germplasm. This will provide a benchmark against which the influences of other breeding approaches and agronomic practices can be compared. Traditional varieties included in the analysis will be obtained from several members of the consortium (partners RIKILT, SCRI, TUM, UKU, IHAR, ISS, CSIR, BiOSS, ICGR). In addition, specific contractors (SCRI, TUM, IHAR, ICGR) have access to traditionally developed breeding lines (diploid and tetraploid) with enhanced "natural" host resistance to specific pests and disease (fungi, virus, bacteria, nematodes). Since natural resistance mechanisms can involve the production of potentially toxic secondary metabolites, these lines will be included in the comparative analyses (for example, the results would or would not indicate these metabolites to be an emerging risk from traditional breeding approaches, providing input for WP2).

1.1.2 Somaclonal variants and lines developed using protoplast fusion*:*
Potato lines developed using protoplast fusion (partner TUM) will be compared with parental wild types. To more fully test the impact of somaclonal variation somaclones of cv Desiree will be developed (by partner SCRI) during the four-year programme to more fully assess unintended effects caused by de-differentiation in callus culture.

1.1.3 GMOs : GMO lines will include (alongside the parental controls) : 1) independent transformation events with enhanced resistance to Potato Virus Y (PVY: partner IHAR); 2) Insect resistant potato (two groups expressing different Bt toxins; partner CSIR); and 3) GM lines with specific quality traits modified (including the levels of nutritional compounds partner SCRI). Links will be established with Chinese Institute for Potato Research thus we envisage the incorporation of further transgenic lines during the programme.

To verify whether the same mold/mycotoxin reduction occurs with insect resistant potato varieties as was observed for GM maize, samples generated from GM-potato will be analysed for the presence of molds/mycotoxins.

Maize

1.1.4 Cultivated maize ("natural" variation).
The objective will be to establish an empirical database using profiling approaches to illustrate existing variation in cultivated maize. Therefore, a broad spectrum of maize varieties will be grown at different locations. In particular, this will include late ripening varieties, which are good candidates to grow in Germany, Italy, and South Africa for comparative purposes. In addition, maize varieties will be bought on the open market and obtained from commercial maize farms. The resulting profiling database will provide a benchmark against which the influences of a number

of breeding approaches and agronomic practices will be compared. Partners will deliver lines from different countries obtained by modern conventional breeding techniques (*i.e. in-vivo* haploid induction). Traditionally bred lines with improved resistance to specific pests and disease, *e.g. Helminthosporium turcicum*, and specific mutant maize, *e.g.* with modified starches, will be included in comparative profiling analyses with special focus on potential unintended effects resulting in the production of toxic secondary metabolites. The exact varieties will also be specified but will at least include non-GM varieties of white and yellow maize. Partners growing and providing maize lines are: TUM, ISS, CSIR, and ICGR.

1.1.5 GM Maize

Insect resistant Bt-maize belongs to the most successful and most frequently planted genetically modified crops. Emerging profiling techniques will be used for comprehensive analytical characterisation of Bt-maize lines which will go far beyond the current state of analyses performed as part of safety assessment frameworks. GM lines (e.g. Mon-810, Bt-176, CIGB-03,MG 95-1, Zhong 31, etc.) and corresponding isogenic lines will be grown at different locations in Bavaria, South Africa and China. Comparative profiling data sets will be assessed using the established profiling database. Later on, comparative profiling analyses will be expanded on additional GM maize lines exhibiting specific traits, *e.g.* glyphosate resistance, enhanced provitamin A content and salt tolerance, etc.

Based on the reduction of mycotoxins on GM maize lines observed for fumonisins and in part also for aflatoxins , the GM lines under investigation in this project will be tested for the presence of these and other mycotoxins (such as aflatoxins, and fusarium toxins) since these may cause health problems for the consumer. The research will be conducted on a sufficient number of samples that can allow for a statistic based conclusion and evaluation.

1.2 Agricultural Production Systems

1.2.1. Potato

Potato varieties (wide range) will be grown under low and high input systems and provided by partners SCRI, TUM, UKU, IHAR, CSIR together with other contractors. Partner IHAR will also provide access to materials maintained in a) a professional store and b) in others such as clamps by taking into account also possible prevention strategies for natural toxicants, fungi occurrence and subsequent mycotoxin formation .

1.2.1. Maize

In order to reveal the influence of different production systems on maize composition and to identify specific risks resulting from these practices, targeted and profiling analyses will be performed using GM and conventional material grown under high and low input regimes in Spain, Germany, South Africa, China and other relevant European Countries.
This comparative analysis will cover commercially produced maize deploying the recently developed prevention strategies (i.e. GAP). Consequently, a full and detailed link with relevant consortium members from the EC, QoL-KA1-mycotoxin-prevention-cluster (contracts n°QLK1-CT-1999-00996; QLK1-CT-1999-01380; etc.) and the relevant EC, INCO-project, SAFEMAIZE (contract ICA4-CT-2000-30033) will be established.

2. Plant Analysis : Profiling Approaches to Differentiate “Risk”

Metabolomics, proteomics, transcriptomics

This study will test the extent of variation in the proteome and transcriptome produced through specific breeding approaches or induced under a range of agronomic practices and environments. Synergies between the various profiling approaches and the utility of such synergies in contributing to the risk assessment process needs to be firmly established. This work programme will provide this. Metabolomics and proteomics analysis will concentrate on both maize and potato, transcriptomics exclusively on potato as the model (potato microarrays are available to the consortium). This could change during the course of the programme to include maize if maize

arrays become easily available within the public domain. Consequently, these changes will then be implemented in close agreement with the Commission. This work programme will make especially use of the results of the GMOCARE project (contract QLK1-CT-1999-00765) and the criteria document published by EC-JRC-ISPRA ("Definition of minimum performance requirements for analytical methods of GMO testing; version 1.7.2003)

2.1 Metabolomics:

The three most important techniques to have emerged are GC, HPLC and NMR and all will at least be used in this project. NMR will be the responsibility of partners RIKILT and CSIR, GCMS and LCMS - principally partners SCRI and TUM. Other contractors will contribute to more traditional analysis, (including some proximate analysis for benchmark comparisons with profiling [e.g. GC, HPLC, etc.]). Profiling methods will be chosen on the basis of their capabilities of detecting, resolving and quantifying a wide range of compounds in a single sample(s). Also a combined approach of GC-TOF-MS and LC-MS^n will be employed, ensuring comprehensive coverage of analytes (non-polar, polar, ionic and protein) and/or others if so necessary and/or required. Measurements for each of the at least three techniques used will be made at least in duplicate for each sample. Early establishment of a scheme to record information on sample origins together with associated files of spectra and chromatograms will be performed. Establishment of optimal conditions of measurement is desirable at the outset, as the aim will be to build consolidated databases as the project proceeds.
Thorough quality control and assurance programme(s) will be implemented within this IP. One primary objective is to optimise the numbers of compounds resolvable by NMR and chromatographic procedures. The use of three complementary techniques will aid the overall task. Data-mining within GC-MS and LC-MS experiments will be aided through the automated use of AMSDIS software which automatically searches for the presence and then quantifies peaks from a user-generated library. Data entry is then direct to Excel and/or other suitable software for various statistical treatments.

2.2 Proteomics

2D gel approaches (IEF; SDS PAGE) will be applied (by partners UKU, CSIR, CIGB and ICGR) with minimal sample processing and gels stained with SYPRO Ruby (protocol validated by partner UKU) which has excellent properties in terms of facilitation of quantitative analysis. Gels will at least be screened with a fluorescent scanner and analysed using PDQuest software to identify qualitative and quantitative differences prior to advances statistical analysis. In addition, to increase the breadth of the analysis. Selected staining protocols will be deployed to analyse the "sub-proteome"– to include post translational modifications (e.g. phosphorylation [signal transduction mechanisms central to plant processes], oxidation [inducible by stress]). Polypeptides which help to differentiate lines bred using different approaches or which respond to high/low input production practices will be subjected to MALDI TOF analysis to assist in product identification. Importantly, pathogenesis related proteins (PRPs) are up regulated under certain disease conditions. PRPs, potential allergens, will at least be assessed in healthy and diseased plant materials. Western blotting will be used to assess potential risk emerging from PRPs in materials recommended by the plant pathologists (see toxins section).

2.3 Transcriptomics

Expression profiling in a robust, accurate, flexible manner will be performed in order to complement and add value to the data obtained from proteomic and metabolic equivalence studies. Partner RIKILT will apply a recently by the Institute developed cDNA microarray with a ca. 4000 tuber unigene set derived from subtractive libraries. Libraries including genes isolated from tubers accumulating natural toxicants such as glycoalkaloids, as well as tuber- and plant-part specific libraries, representing the edible and non-edible potato plant parts will also be used. The plant-part specific library represents for a large part metabolic routes that are usually silenced in the tuber and may therefore give insight into possible shifts in activated pathways in the tubers under investigation. All clones have been sequenced and selected on the basis of the resulting data and a subsequent comparative analysis with the NCBI database. The microarray approach will be used to assess modified gene expression in pathways related to nutritional and anti-

nutritional factors in the potato and importantly will provide the first benchmark of variation in gene expression in mature tubers from a range of genetic and environmental backgrounds. To this end additional DNA probes from metabolic routes of interest will be obtained on the basis of published sequences or will be selected from the available libraries and/or purchased from other sources, cloned and included in the array. The aim is to have all key metabolic routes represented on the array to gain optimal insight into possible alterations in the plant's physiology. The array occupies *c.* 1 cm^2 and as the amount of mRNA necessary for a single hybridisation is dependent on the surface of the array, this reduced size will make it possible to hybridise a significant number of potato samples with a limited number of replications for the screening of variations in gene expression. For construction of the microarrays a suitable pipetting robot(s) and a microgrid II array spotter will be used for optimal accuracy and high throughput manufacturing. The project aims to increase the sample throughput by optimising especially the RNA isolation procedure. For detected differences in gene expression that may have relevance for the food safety of the plant product additional tests will be carried out to further investigate the variation in gene expression in the (GM) varieties of interest. To this end real-time PCR methods will be used and/or developed on the basis of available sequence information, utilising the lightcycler system. Selection of the real-time PCR methods to be set up will be done on the basis of the results of the different analytical methodologies in the individual (GM) lines in the project. In close cooperation with the BIOSS bioinformatics group (partner BiOSS), a data system will be initiated that includes all relevant data on the potato lines under investigation, the experimental set-up, data handling and analysis. Gene expression analysis of novel potato lines will be done on the basis of the data obtained from the range of conventional potato varieties that are already on the market. To this end potential differences in gene expression will be identified based on selected confidence intervals per spotted sequence resulting from the accumulating data on conventional potato varieties. This data system will serve as a extensible background database for future experiments to assess altered gene expression in tuber samples.

3. Emerging health risks arising from plant pathogens

Rationale: Fungal Pathogens and Mycotoxins
This project will make active use of the available and appropriate methodologies in order to identify and quantify the risk for human and animal health associated with the presence of fungi and mycotoxins. Methodologies include characterisation of toxigenic fungi and identification of their toxic metabolites with respect to microbiological contamination by also fully adhering to the minimum acceptance criteria obtained and published by the EC, QoL-KA1-mycotoxin-prevention-cluster. For mycotoxins sampling procedures (i.e. the approach specified by the EU regulation and / or respecting a CV $\leq$ 10%), analytical methods and reliable monitoring studies have been developed and will be applied, which are based on the criteria specified in CEN & ISO and / or other documents from relevant EU regulatory project outputs, the PREMYTOX – project, and the COST 835 – action on mycotoxins. This component of WP1 will consider potato only for eventual development of novel analytical procedures and novel mycotoxins.

Maize

The analysis of maize will include determinations of the major mycotoxins (including *fusarium* toxins) present in GM and non GM lines (different varieties, different agricultural practices and storage conditions) by also taking into account the recently developed and published prevention strategies from the mycotoxin-prevention-cluster and the COST 835 action and PREMYTOX. This will require the application of current sampling methodologies by all partners. In addition, fungal populations in South African maize will be enumerated and compared to identify those toxigenic fungi for which the mycotoxins are known and unknown (e.g. *Stenocarpella maydis* and *Phoma sorghina*). This will give an overall comparative view of the risks arising from any introduction of new fungi in GM and non-GM lines. Here, active use of the recent results of the EC, INCO-project on SAFEMAIZE or other relevant projects will be made.

Potato
New mycotoxins and/or natural toxicants will be identified and methods will be developed for the unknown ones. Their occurrence in potato from different agricultural practices deriving from other partners will be evaluated. Researches will also be carried out on commercially available derivatives in order to investigate the fate of the toxins in intermediate and finished products. This will require the application of current sampling methodologies as laid down in European regulation and/ or new ones by all partners and respecting a maximum CV $\leq$. 10%.

Experimental Approaches:

3.1 Enumerate and identify fungal/bacterial pathogens: from both maize and potato (transgenic and non-transgenic, high and low input systems; comparative storage regimes [partners IHAR, ISS, CSIR and ICGR]). Bacterial pathogens are included where appropriate to health e.g. *E coli* and total coliforms, enterobacteriaceae, *Salmonella*, *Staphylococcus aureus, Clostridium perfringens*, *Bacillus cereus*, and *Listeria*. Total counts of aerobic and anaerobic mesophilic organisms will also be carried out. PCR analysis will be used to complement traditional mycology/bacteriology approaches and use will be made of the methods and approaches from the following EC-supported projects: Bacillus Cereus (contract QLK1-CT-2001-00854), Demowatercoli (QLK1-CT-2001-01209) and BACANOVA (QLK1-CT-2001-01145).

3.2 Mycotoxins: will be performed with HPLC, GC, and/or other relevant techniques (partners ISS, and ICGR) all of which will fully adhere to the different acceptance criteria specified by the Commission, CEN, ISO and those generated by the linked mycotoxin prevention cluster and Codex. In cases where previous detailed analytical studies have failed to reveal the nature of the mycotoxin produced by the toxigenic fungus, it will be necessary to rely on fungal detection and enumeration.

4. Statistics and Data Analysis

Partner BiOSS will centralise and drive the key components of comparative statistical analysis to ensure uniformity and transferability of data outputs into the development of risk models. Data outputs will be co-ordinated in the most appropriate form. An important consideration will be to maintain consistency in extraction and measurement procedures between analyses that may be separated by several months (since it is intended to construct consolidated databases). Use of internal standards and references will allow to compensate for changes in machine performance, columns etc, as well as to fully adhere to the different acceptance criteria among those specified by the Commission, JRC-ISPRA on GMOs, CEN, ISO. Validation of the protocols and analytical methods to be used will be performed at the very start of the project with guidance of professional statisticians and by using certified reference materials wherever possible.

4.1 Statistical Approaches
Two approaches will be taken to statistical analysis of the data. The traditional univariate approach (analysis of variance, REML) will be applied to establish significant differences between positive and control group mean values, or quantify the levels of random variation from theses sources, on a compound-by-compound or peak-by-peak. In the second approach multivariate analysis methods will be applied, e.g. to GC-MS outputs and for sample classification. The data to be analysed may be in a “raw” form or as tables of absolute or relative amounts of specific compounds. Principal component analysis (PCA) will be the main method of data compression. It reduces the original data (intensities for hundreds or data points or metabolites for each sample) to a set of “scores” on a much smaller number of principal component (PC) axes for each sample. The scores are first used in an exploratory way to look for clusters of similar samples (e.g via scatter plots) and then in a more formal fashion to construct classification models.

Once the original spectra are compressed to PCA scores, cross correlations will be assessed. Multivariate methods such as principal component regression and partial least squares will be used

along with simple matrix scatter plots. Establishing such links will improve our understanding the biochemical relationships and will provide a more comprehensive picture of the substantial equivalence between the GM and non GM lines and between samples derived from high and low input systems.

4.2 Interpretation and models of unintended effects.

Assuming a sample or group separation is established using profiling data, the univariate, peak-by-peak, approach leads directly to identification of the peaks responsible and these can be assigned to specific metabolites if assignments have been possible for the relevant peaks. On the multivariate approach similar information will be obtained from the PC loadings. Beyond this identification, interpretation of the underlying mechanisms behind the changes will have to draw on existing knowledge of metabolic pathways. However information will also be uncovered (concerted increase of certain compounds, increase of one group of compounds accompanied by decrease in another group), that will aid new interpretations, especially when the intended modification is precisely known.

An attempt will be made to show how new individual samples might be tested for acceptance/rejection with respect to an accumulated database of ‘acceptable’ samples, somewhat akin to the problem of medical diagnosis (highlighting abnormal samples against a background of ‘healthy’ ones, even though the definition of ‘healthy’ covers an enormous range). Methods based on PCA such as SIMCA (soft independent modelling of class analogy) will be applied to deal with this type of problem but requires a large number of samples to model the ‘acceptable’ class. Therefore this task will be left for the final stage of the project when a large database has been accumulated for each of the techniques. An evaluation and comparison of the different spectroscopic/ chromatographic techniques will be made with a view to recommending the most suitable procedure for practical adoption. This comparison will be aided by the fact that a common set of samples will have been measured by all three techniques. Measurement of some samples by the same technique on two or more different sites will aid the development of the techniques, particularly regarding the need to achieve standardisation and consensus.
The comparative profile database used by risk assessors will be further assessed with respect to their validity. Identifying potential differences between foods produced in the different systems, which may be of health significance, against a background of natural variations, is of great help to develop a more balanced and science-based risk assessment strategy for foods produced in different systems. The profiling approach and generated information will be used in Research work packages 4, 5 and 6 in order to design strategies for risk assessment policies and for communication with consumers.