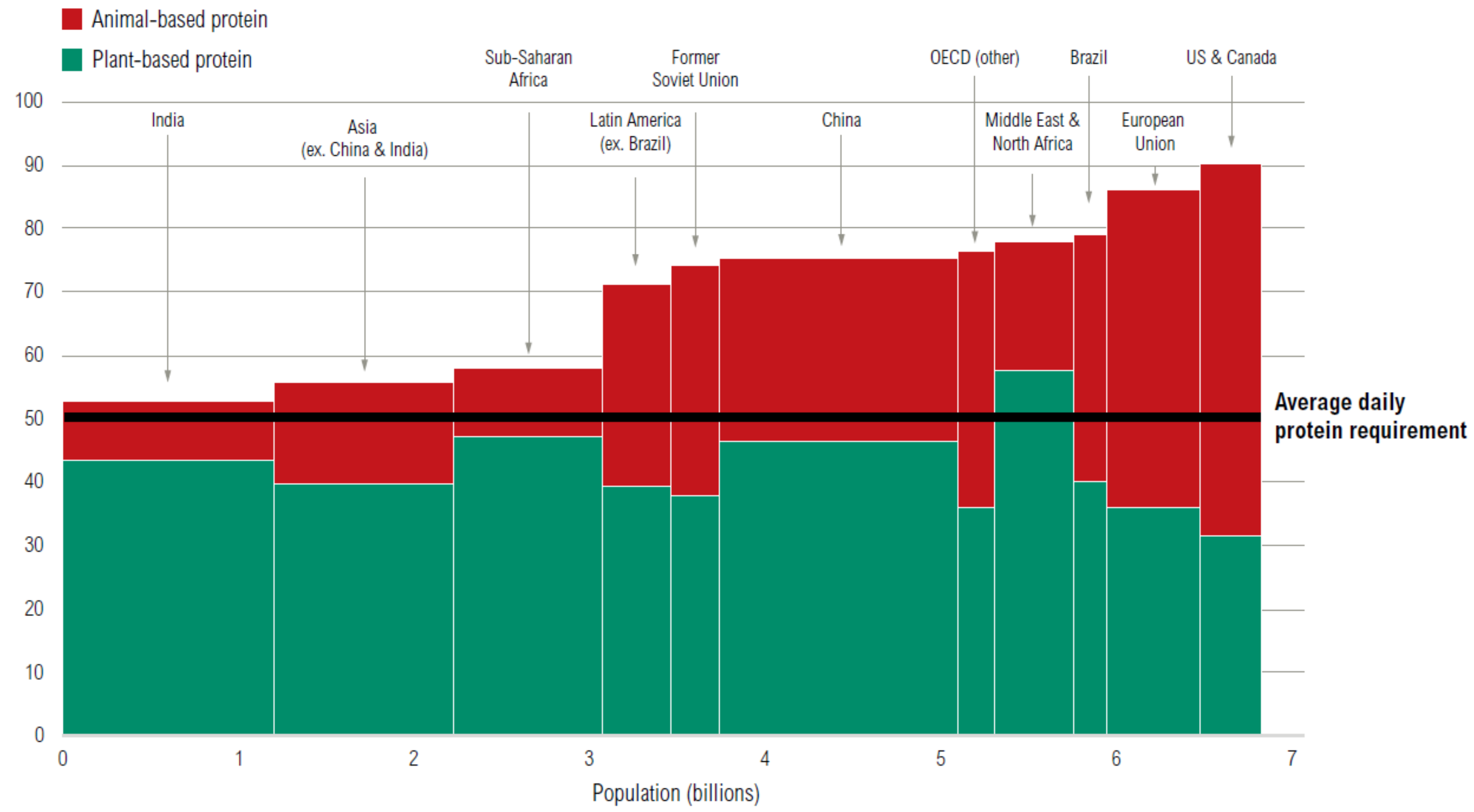
THIS PROJECT IS FUNDED BY THE **EUROPEAN UNION HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME** UNDER GRANT AGREEMENT 817690

N-1: Increasing protein content and quality

Agata Gadaleta[1], Benoit Méléard[2], Bertrand Hirel[3], Carlos Guzmán[4], Catherine Ravel[5], Friedrich Longin[6], Gilberto Igrejas[7], Jacques Le Gouis[5], Karine Gallardo[8], Peter Shewry[9]

1: University of Bari "Aldo Moro", Bari, Italy
2: Arvalis, Boigneville, France
3: CNRS, Versailles, France
4: Universidad de Córdoba, Córdoba, Spain
5: INRAE, Clermont-Ferrand, France
6: University of Hohenheim, Stuttgart, Germany
7: University of Trás-os-Montes and Alto Douro, Vila Real, Portugal
8: INRAE, Dijon, France
9: Rothamsted Research, UK

Protein consumption around the world



Source: GlobAgri model with source data from FAO (2015) and FAO (2011a). Width of bars is proportional to each region's population. Average daily protein requirement of 50 g/day is based on an average adult body weight of 62 kg (Walpole et al. 2012) and recommended protein intake of 0.8 g/kg body weight/day (Paul 1989). Individuals' energy requirements vary depending on age, gender, height, weight, pregnancy/lactation, and level of physical activity.

Ranganathan et al. 2016

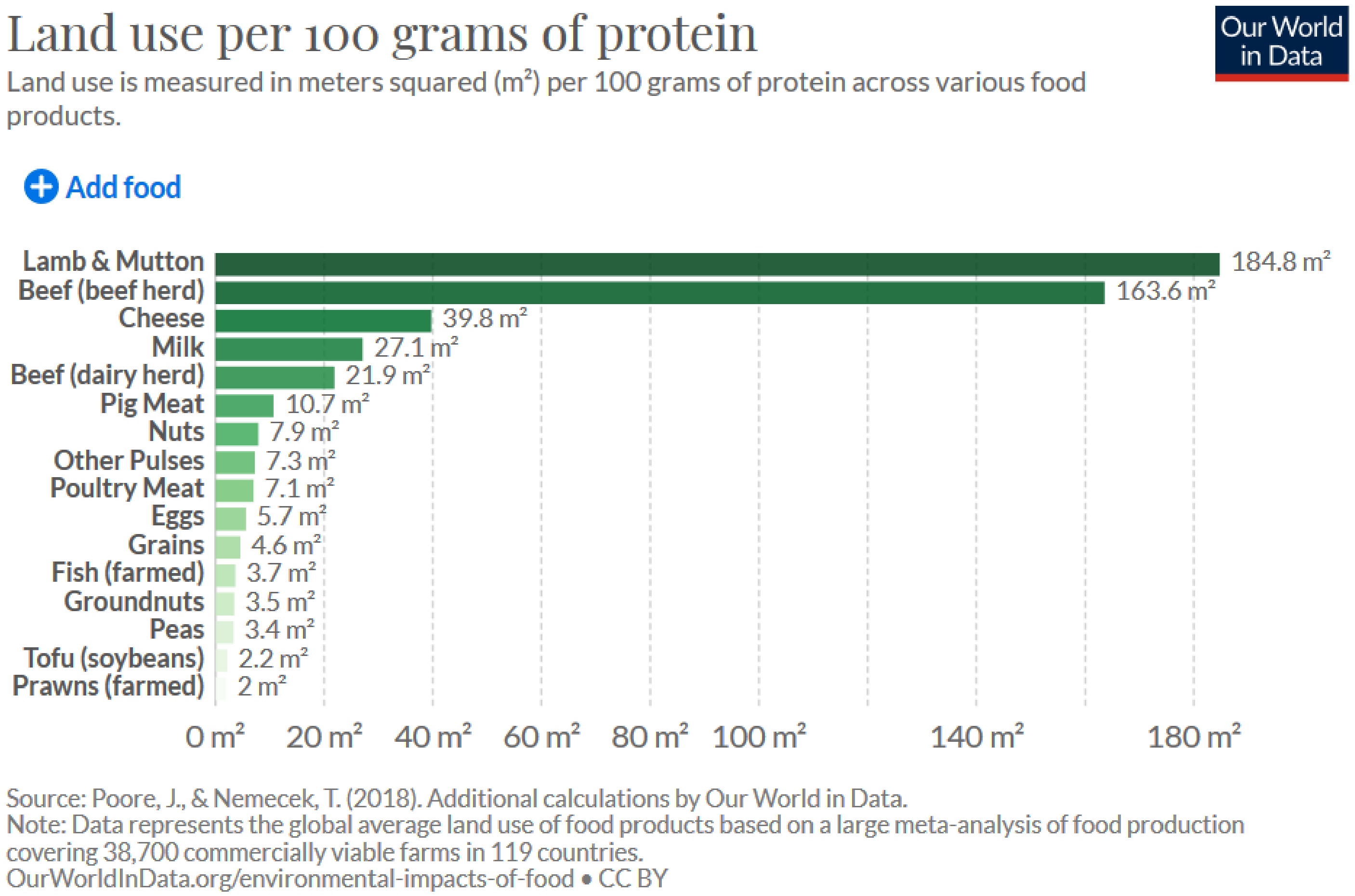
Shifting to a vegetable diet

Animal-derived protein consumption has increased

They are highy digestible proteins well balanced in aminoacids

But associated issues are:

- Greenhouse gases emissions
- Health related problems
- Terrestrial biodiversity loss



Sustainable plant proteins production

EU strategic goals, set and documented in the "Farm to Fork" agenda, look forward to:

- **Increasing plant proteins contribution to the diet**

"A key area of research will relate to microbiome, food from the oceans, urban food systems, as well as increasing the availability and source of alternative proteins such as plant, microbial, marine and insect-based proteins and meat substitutes;"

- **Fostering UE-grown proteins**

"The Commission will examine EU rules to reduce the dependency on critical feed materials (e.g. soya grown on deforested land) by fostering EU-grown plant proteins"

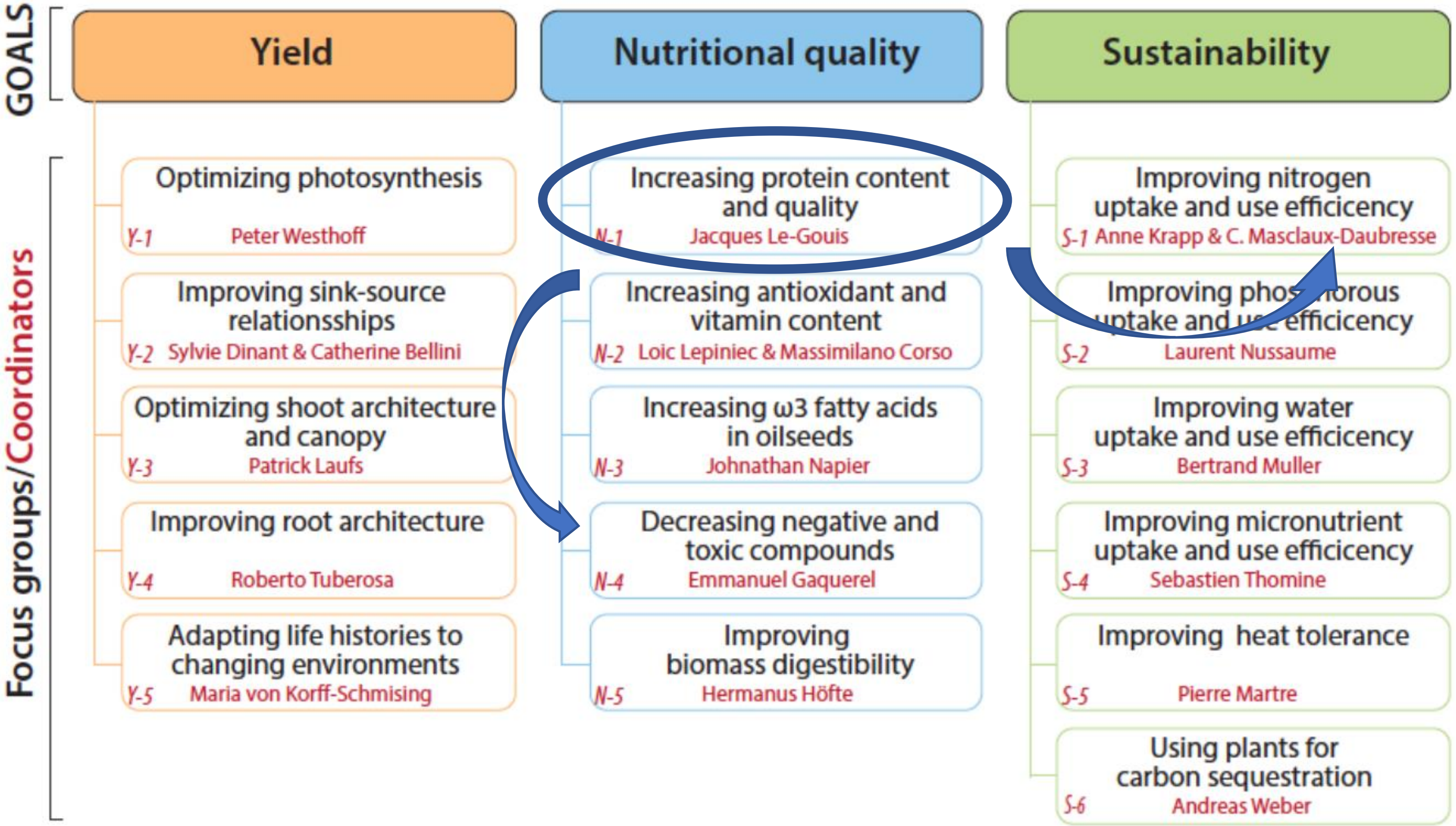
- **Reducing the use of chemicals**

The goal is to "Reduce significantly the use and risk of chemical pesticides, as well as the use of fertilisers […]" and to "Increase the area under organic farming in Europe."

- **Adapting the crops to climate change**

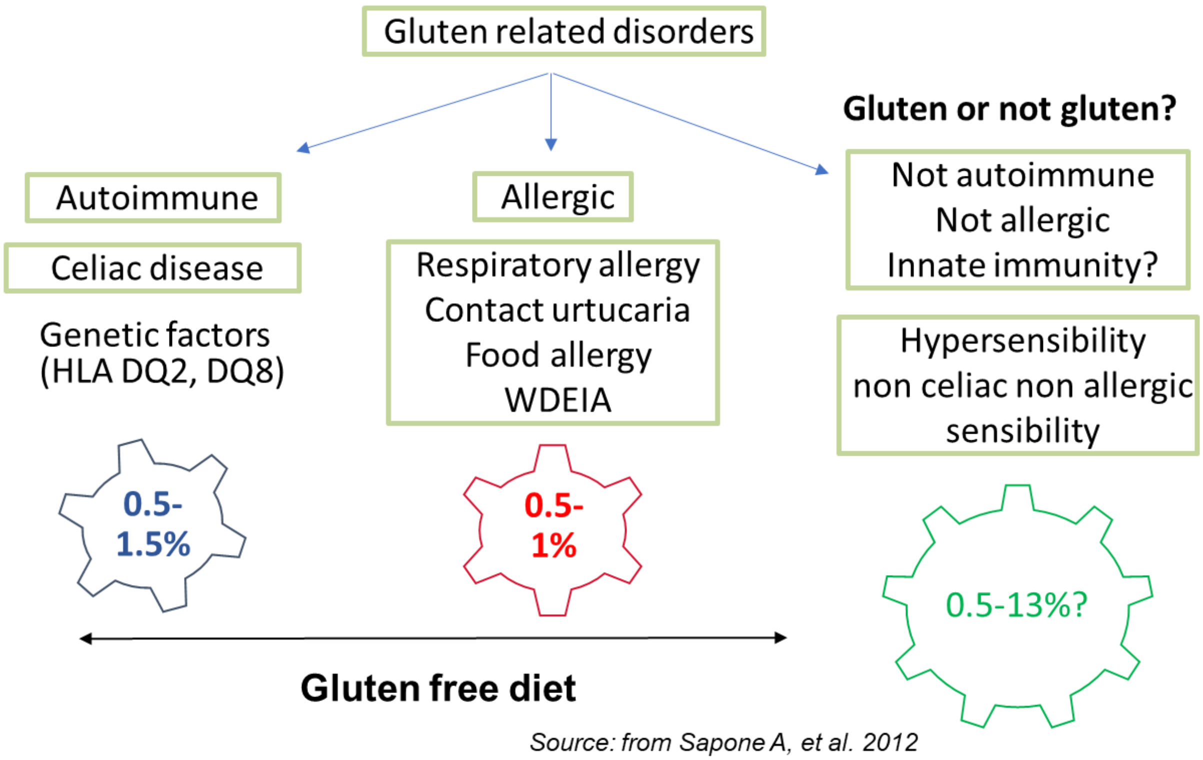
Focus of the group

"Increasing protein content and quality" is quite large so it was proposed **to really focus on the content and composition of the grain proteins for technological and nutritional quality**



Protein negative effects

Gluten related issues in *triticeae*



Antinutritional factors for legumes

- trypsin inhibitors, tannins, vicin & convicin for faba bean have been reduced
- need to reduce allergens and improve organoleptic properties (reduce the bitter taste in part due to saponines)

Focus on three main crops

Small grains cereals (Wheat)

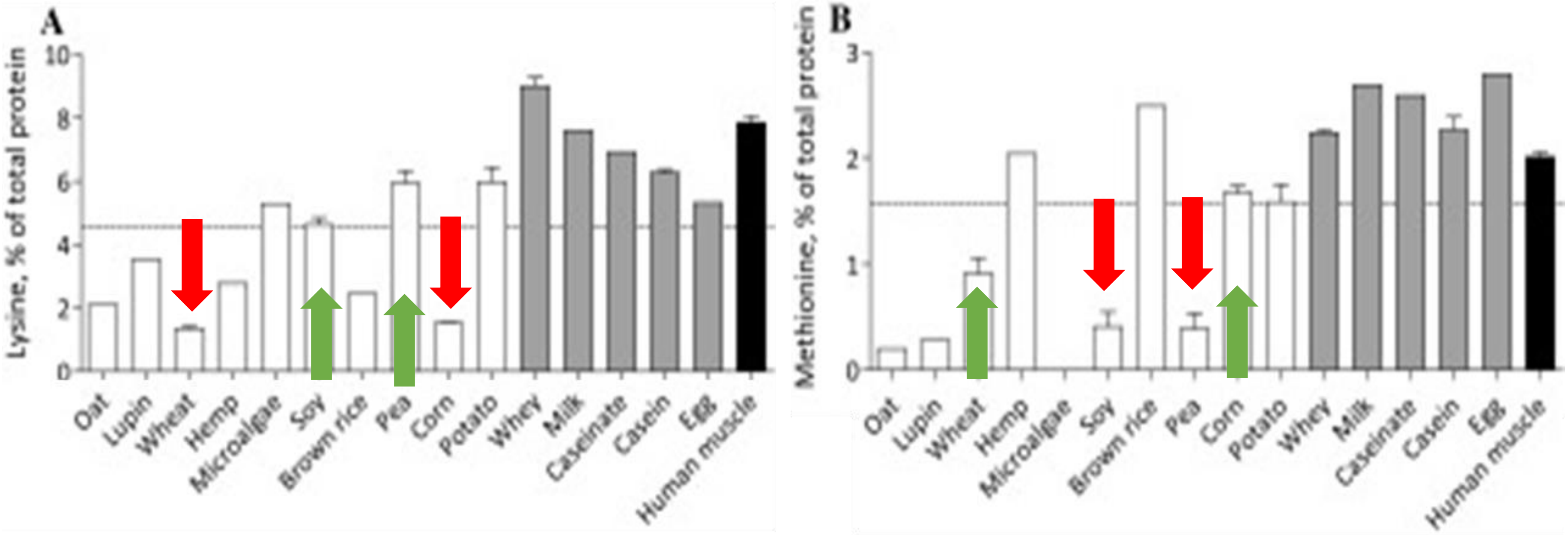
- C3
- feed and food

Maize

- C4
- mainly feed

Pulses

- C3, N symbiotic fixation
- mainly feed



Gorissen et al. (2018)

Current know-how : major genes known

- Variation for processing and end-use quality largely **explained by protein content and composition**

- Major loci or genes for both protein content and composition have been cloned

7S and 11S globulins in legume seeds, glutenins and gliadins in wheat, zeins in maize

- **Some regulatory genes have been identified**:

modifiers of the opaque-2 mutation in maize

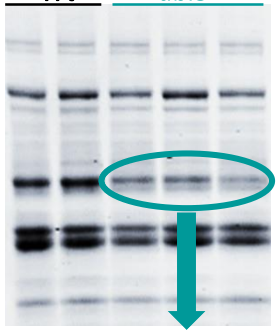
ABI5 for vicilin (7S) synthesis in pea seeds

A PQL (protein quantity loci) analysis of mature pea seed proteins identifies loci determining seed protein composition

Michael Bourgeois[1], Françoise Jacquin[1], Florence Cassecuelle[1], Vincent Savois[1], Maya Belghazi[2], Grégoire Aubert[1], Laurence Quillien[1], Myriam Huart[1], Pascal Marget[2] and Judith Burstin[1]

Genome-wide association studies with proteomics data reveal genes important for synthesis, transport and packaging of globulins in legume seeds

Christine Le Signor[1], Delphine Aimé[1], Amandine Bordat[2], Maya Belghazi[3], Valérie Labas[4], Jérôme Gouzy[5], Nevin D. Young[6], Jean-Marie Prosperi[7], Olivier Leprince[8], Richard D. Thompson[1], Julia Buitink[8], Judith Burstin[1] and Karine Gallardo[1]

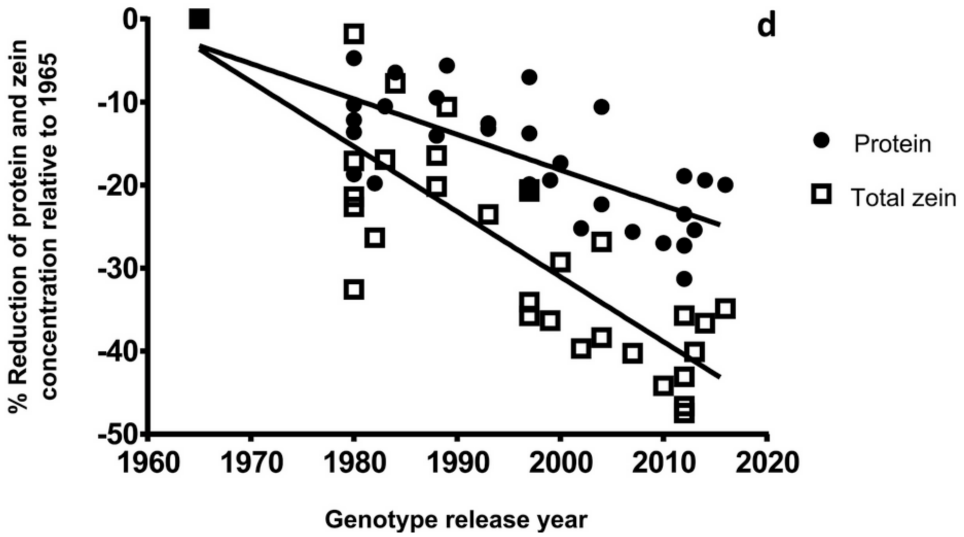


abi5 pea TILLING lines poor in vicilins

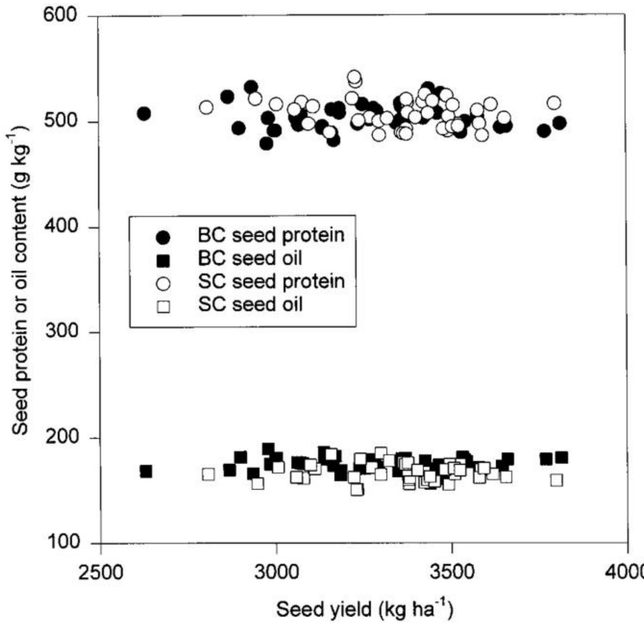
- **large genetic variability still exists for protein content and composition**
  - for end-use quality (wheat, pea)
  - for nutrition quality (wheat, maize), e.g. sources rich in Lysine

Current know-how : negative correlations to yield

- Whatever the species, trend is observed for **lower grain protein concentration for modern cultivars**
  - maize hybrids released in Argentina, Caballero-Rothar et al. (2019)

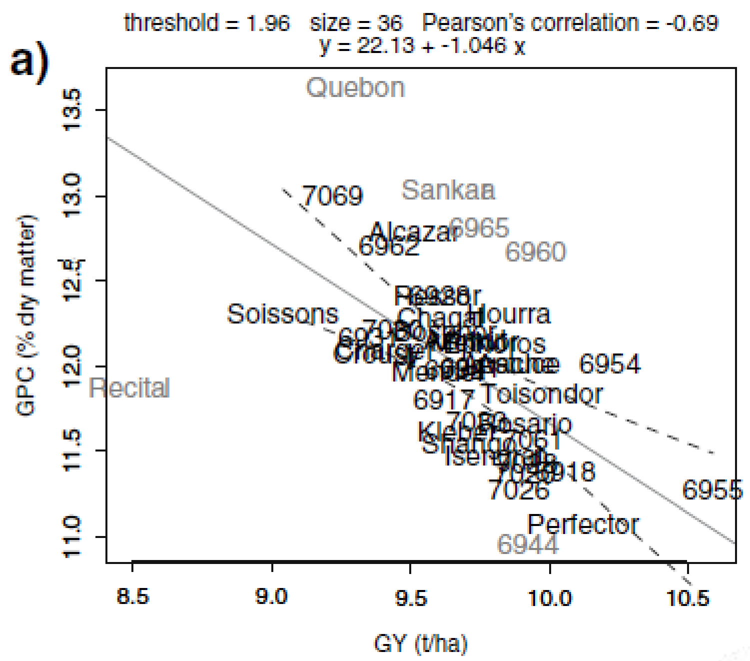


- However, **possibility of improving the yield while maintaining seed protein content** in grain legumes and possibly maize, as no strong negative correlation was observed
  - soybeans, Cober and Voldeng (2000)

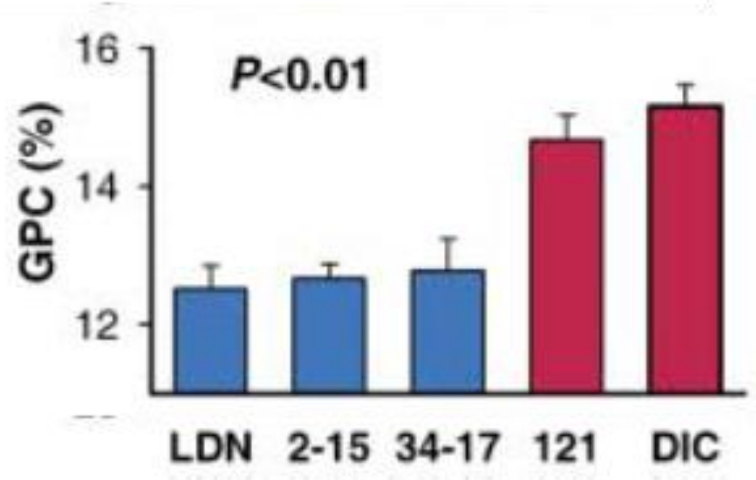


Current know-how : use of the GPD criteria in wheat

- In cereals, exploitation of the **deviation from the negative relationship between grain yield and protein concentration (GPD)** possible as shows its stability and heritability
  - Oury et Godin 2007

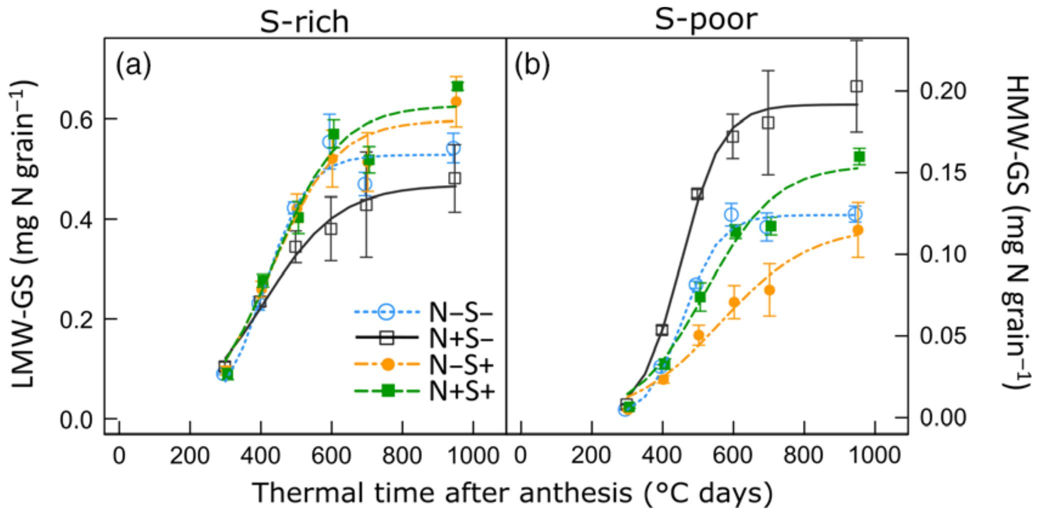


- Some regulators, as the **TaNAM transcriptional factor** in wheat, gives hope of being able to increase the protein content without negatively affecting yield
  - Uauy et al 2006 (Langdon x T. diccocoides)



Current know-how : strong impact of environmental factors

- In addition to genetic factors, protein content and composition are **strongly affected by environmental factors** : pedo-climatic (soil properties, temperature, precipitation, CO_2, etc.) and crop management practices (e. g N/S fertilisation management, crop protection …)
    - Bonnot et al., 2017 on wheat



- Effects of low nutrient (N, S) availability on seed protein content and composition have been studied, notably in combination with drought, and **candidate genes that might limit the negative impacts of these stresses** on seed development were identified
    - Henriet et al., 2019 & 2021 on pea

An antioxidant protein network underlies the response of developing pea seeds to sulfur deficiency and water stress

Trends in research

- **sustainable field managements** highly investigated due to the development of low input systems and climate change

- effect of **modern breeding on grain quality**
  - cereals : increase of wheat related health disorders
  - legumes : negative effects in animal feed eliminated by selection although they can have positive effects in human health (e.g. tannins with antioxidant functions)

- **more and more investigation of combinations instead of individual factors**: % of glutenins/gliadins/albumins/globulins; glutenins alleles combinations for the wheat grain, % of 7S globulins/11S globulins/2S albumins in legume seeds

- it is recognized that there is a need to understand effects at all levels: **from gene expression to the phenotype and the biophysical effects on molecular interactions**

- **translational genomics** is proposed when relevant to take advantage of information coming from different species

Relevant technologies

- **Genomes of all the major crops have been sequenced**
    - Very good reference genomes
    - More and more sequences are now available to describe the pangenome
    - High density genotyping arrays and genotype by sequencing developed for GWAS and GS

- low cost and high throughput **proteomic tools,** particularly adapted to analyzing grain proteins related issues

- **TILLING populations (**EMS-induced allelic variations) developed (in maize, wheat and pea) : particularly efficient for traits determined by major genes

- **genome editing**, that can target multicopy genes (polyploid species like bread wheat) available

- **multi-level data integration tools** (phenomics, proteomics, transcriptomics, genomics…)

- **ecophysiological modelling** could be a relevant tool to understand the interplay between environmental and agronomic factors

Future challenges : diversity and stability

- The quality of a production **is a very complex and changing concept**. Necessary to understand better in each case (processing, end-use, nutrition health) what are the required criteria and possibly **diversify and specialize the varieties**

- Both a **high quality** (whatever the criterion) and a **stable quality** are required. In that context the effect of global change is a major challenge (low input systems, drought and heat stresses; elevated CO_2; etc.) for breeding

Future challenges : maintain quality

- Better understand **correlation between grain yield and grain protein concentration**

- Whatever the species there is need to better understand this correlation and possibly identify genes that can alter it (notion of GPD in wheat)
- For cereals, there is then a need to improve the efficiency of conversion of applied N to grain protein
- Other **trade-offs may exist**, for example the consequences on resistance to biotic & abiotic stresses (cold, drought, fusarium & bruchids in legumes ...) of the elimination by genetics of antinutritional / off-flavor factors

- Assess how it is possible to keep the processing and end-use quality levels **when protein content is going down** (because higher grain yields, less fertilization or less efficient N fixation) by improving the protein quality and **finely tuned the protein composition**

- Develop **socioeconomic analyses** along crop rotations regarding nutritional proteins, e.g. is it necessary to have wheat with 12% protein when legumes can deliver much more. Studying societal acceptance of the consumption of legumes is needed to develop their production for human nutrition.

Future challenges : tools developments

-The quality of a product is estimated generally only on a few easy and rapid to measure criteria (e.g. grain protein concentration for wheat). **Development of rapid tests** are needed (for baking quality so that breeding and trading for real baking quality and nutritional aspects for humans and animals can replace simpler tests)

- A technological/scientific challenge is the establishment of **good data bases** for the exploitation of new proteomic tools, as most of the identified proteins are unknown

- One main societal challenge is probably around the possible **acceptance of genetic engineering (GMO) and CRISPR modified crops** (targeting protein quality) for commercial use in the long term if positive benefits on environment and health are demonstrated

Opportunities for genetics and breeding

- In the context of global change, along better management strategies, **opportunities to improve and stabilize the quality with reduced levels of inputs**, use of fertiliser (lower costs and environmental footprint), pesticides and energy required for processing

- In the context of an increase of gluten related disorders, **opportunities to propose innovative wheats which solve gluten sensitivities**

- **Marker/genomic assisted breeding** is fast developing to combine best genetics for yield-protein content-protein quality. Many QTL were previously described but not often used in real breeding programs.

- **Use of translational genomics** to transfer knowledge of the regulation of seed protein content and quality from models to crops (e.g. *M. truncatula* or *Arabidopsis* to legume crops) or between crops (e.g. pea-faba bean-soybean-lentil…) is increasing

- Development **of predictive models (ecophysiology x genetics) to estimate grain N content**, N translocated to the grain (grain N removal) and N balance in the soil taking into account environmental factors (temperature, water variability) and considering genetic parameters linked to the genotype

Action points

- **to exploit the genetic diversity** (natural as well as induced by TILLING) to increase the quality through improved pre-breeding / breeding is still a challenge

- still the need to **better understand the genetics of protein content** (nitrogen use efficiency / interaction with microbiota) to breed varieties competitive for grain yield, with same or more protein content and that demand less fertilization

    - This includes identifying the molecular determinants of components responsible for off-flavours of legume-derived ingredients, and of for gluten-related diseases

- greatest challenge is **to understand G x E x M interactions** for seed protein content and composition.

    - Should include climate change and agroecological practices (e.g. the use of legumes in cereal cropping systems)
    - Should include biophysical effects as well as effects on gene expression

Action points

- **breeding and registering new varieties is a way to translate** research into societal and economic value
    - Translational mechanisms to deliver traits and markers to breeders may be also through pre-breeding materials (e.g. UK BBSRC Designing Future Wheat programme).

- **Propose new quality criteria that could integrate the registration system**
    - GPD is a registration criterion assessed in the French registration system for Bread Wheat

- More **projects/collaborations between academia-breeding companies-food industry-extension services-farmers-consumers** are required

- More presentations of the challenges, projects and results regularly on scientific but also **applied meetings with stakeholders and consumers using Blogs and social media**

- Developing **attractive plant protein-based products with quality labels**
    - Better balance the intake of proteins from animal and plant sources. Complementarity of different sources of proteins (cereals, legumes, Brassicaceae) to meet the requirements for essential amino acids

Genetic levers to improve quality

Exploit genetic diversity

Develop attractive plant protein based products
Identification of quality criteria and rapid efficient tests

What is the variability in the composition in storage protein fraction?

What is the composition in aminoacids?

What is the genetic control of the composition?

Optimize **Stabilize**

Protein content and composition

Diversify

What is the stability of the concentration in response to abiotic and biotic factors?

What is the correlation to yield or trade-off with other traits?

What is the genetic control of the stability?

Candidate loci and genes for breeding

Combination of genetics, ecophysiology & omics approaches
Combination of protein alleles / traits
Modeling to analyze GxExM interactions
Translational genomics

Thanks for your attention