The main aims of the planned research were the allergen characterization of the seed proteins of the einkorn collection of the Cereal Gene Bank of our Institute in Martonvásár using bioinformatics, proteomics, immunomics analyses. Use of these data obtained from the comprehensive computational data analyses could result in a more effective prediction of the allergen capacity of different wheat genotypes. The main goal of this project is the epitope mapping and expression levels in the different einkorn genotypes at a multidisciplinary level. The large scale immune analytic screen and toxic protein detection of the approximately 200 different einkorn genotypes with human sera serve as important information for clinicians, plant scientist and for the food industry. Based on the results, the further aim of the project had been the detection einkorn genotypes could contribute to the breeding of low allergen wheat genotypes.

The project started with extensive bioinformatics analyses of the *Triticum monococcum* ssp. *monococcum* and *T. monococcum* ssp. *aegilopoides*. This workflow involved the epitope mapping of einkorn and wild einkorn storage proteins based on the publicly available sequence and epitope data. From UniProt database all of the celiac disease related storage protein type of einkorn and wild einkorn were collected. The analyses included alpha, gamma, and omega gliadins, HMW- and LMW glutenins, and avenin-like storage proteins. T-cell and B-cell epitope predictions were made, the Immune Epitope Database (IEDB) (http://immuneepitope.org) was used to retrieve epitopes identified from *Pooideae.* Frequency and distribution of epitopes were calculated both in undigested and digested proteins (pepsin, trypsin, chymotrypsin), and epitopes resistant to enzymatic cleavage were identified. The results of this in silico study was, that einkorn storage protein sequences also contains digestion resistant epitopes, but less, than haxaploid wheat, but the toxic core epitope regions unfortunately are present in most of the investigated sequences. The lack of the core epitope regions or digestibility of them in case of some gliadins presented in the Uniprot database was very hopeful.

The results of our recent research about the epitope mapping and serological analyses of the non-prolamin storage proteins of *Brachypodium distachyon* gave useful information for the investigation of diploid einkorn non-prolamin type storage proteins (Gell et al., 2016).

In parallel with the bioinformatics study the total protein extraction and SDS-PAGE of the einkorn gliadin were made, the protein profiles from 6-6 grain were analysed, and unfortunately during this experiment inhomogeneous genotypes were found, so individual plants had to grow under greenhouse conditions. From the isolated total storage proteins the custom developed ELISA tests were optimized for the investigation three different wheat related disorders (celiac disease CD, wheat allergy, WA and Crohn's disease) and measure the toxic behaviours the einkorn’s proteins. There were different tasks we had to optimize, like the coating protein amount, the dilution of the sera and dilution of different secondary antibodies. After this optimization the large scale custom developed ELISA tests were started with the protein extracts. In this study two negative, three positive on strict diet per disease and ten positive, diet native sera per disease were used. PWG gliadin (10 ppm like in case einkorns) and Chinese Spring proteins extract were used as ELISA standards. Proteins of the investigated genotypes were also analysed with custom developed ELISA tests with and without digestion using trypsin. During the optimization procedure, in case of wheat allergy and specific IgE response we have found that

the sera from wheat allergy affected patients have to use more concentrated as we imagined and due to this fact in the next year in case of allergy, the 200 different genotypes were not able to analyse with the same sera, because of the limitations of its amount.

In the second year the analysis of the total protein extracts of the 200 einkorn genotypes (provided by the Cereal Gene Bank of the Department of Plant Genetic Resources and Organic Breeding) with custom developed ELISA tests were continued. The investigation included three different wheat related disorders (celiac disease- CD, wheat allergy- WA and Crohn's disease) and measure the toxic behaviours the einkorn's proteins. In this study two negative, 3 positive on strict diet per disease and 5 positive, diet native sera per disease used, but in case of celiac disease more positive sera were needed (10) because of the variability the genetic background of this disease (DQ8/DQ8, DQ2/DQ2, DQ2/DQ8 alleles). PWG gliadin and Chinese Spring proteins extract were used as ELISA standards. Proteins of the investigated genotypes were also analysed with custom developed ELISA tests with and without digestion using trypsin. During the genotype selection procedure individual sensitivities of patients were also fund, but only those genotypes were taken into account, that's gave with all of the positive sera lower immune reactive signals.

Unfortunately because of the large sample set, the usually used methods like gel-filtration or size exclusion chromatography could not applied for the decontamination procedure of the trypsin and chymotrypsin residues in the samples. Without this step the digested samples gave higher absorbance in the ELISA tests because they interfere somehow with the investigated peptides. So as an alternative way different pore sized Millipore columns are using (30, 50, 100 kDa) now to determine the toxic behaviour of the digested samples. In case of the lower allergen- or toxic protein containing einkorn genotypes Osborne fractionating were carried out, and the three different storage protein fractions, like water-saline soluble albumins/globulins, the alcohol soluble gliadins and the glutenins further ELISA analyses were made with 5 different positive human sera. In case of wheat allergy the results are not surprising, that the immune dominant part of the seed storage proteins are the albumins and globulins, but concerning the celiac disease it is very important, that the albumin/globulin type proteins gave very high immune reactions with positive sera, provides the following evidence, that these proteins of cereals act as cross-allergens with the trigger gliadins because of the partial sequence homology.

The next major step in the second year was the proteomic analysis of the selected genotypes. Initially, the cleaning of the total protein extractions was optimised for 2D-GE and for the immunoblot purposes (acetone precipitation, desalting, 2D clean up, the IEF profile and the composition of the IEF buffer). After the optimization, isoelectric focusing and 2D separation of the selected genotypes was carried out.

RP-HPLC and MALDI-TOF analyses were made of the selected einkorn samples, these experiments were not included in the originally research plan, but because of the limitation of human sera (for immunoblot 100 microliter per minigel is needed) we certainly wanted to be sure in the correct/appropriate sample selection per disease.

The six selected einkorn genotypes are overlap with each other per wheat related diseases; usually the toxic protein contents reach the 20-25% of the control hexaploid Chinese Spring genotype.

In the third year the analysis of the total protein extracts of the selected einkorn genotypes with custom developed ELISA tests were continued. The investigation included three different wheat related disorders (celiac disease- CD, wheat allergy- WA and Crohn's disease) and measure the toxic behaviours the einkorn's proteins. Proteins of the investigated genotypes were also analysed with custom developed ELISA tests with and without digestion using trypsin. During the genotype selection procedure individual sensitivities of patients were also fund, but only those genotypes were taken into account, that's gave with all of the positive sera lower immune reactive signals. In the last year an interesting phenomenon was detected concerning the protein digests, the usually used methods like gel-filtration or size exclusion chromatography could not applied for the decontamination procedure of the trypsin and chymotrypsin residues in the samples. Without this step the digested samples gave higher absorbance in the ELISA tests because they interfere somehow with the investigated peptides. So as an alternative way different pore sized Millipore columns are using (30, 50, 100 kDa) now to determine the toxic behaviour of the digested samples. In this the problem solved, used a special Proteo-Spin One Column Digest kit. The results of the serological ELISA test with the digested seed storage protein samples were evaluated, and based on it the einkorn genotypes have more digestion resistant storage proteins like modern bread wheats. It is possible, that proteins derivate from the A genome are same behaviour in hexaploid wheat. While after the peptic-tryptic digestion the immune reactive signals of the Chinese Spring control bread wheat decreased to 50%, the einkorn genotypes in average decreased only with 30%. The immunogenic analyses of the investigated genotypes were carried out with commercial available R5 and G12 antibodies, these analyses were not included in the original work plan, but it is very surprising, that despite the serological and immunoblot positive results, the commercial available test kits using for food gluten free diagnostic, not provided this positive immune reactions of the selected einkorn genotypes. Based on the results of R5 and G12 ELISA evaluation these genotypes are gluten free and edible for patients suffering in celiac disease or wheat allergy. The reliability of these test in some cases (investigated and proved in our previous research) is questionable; the used antibodies not recognize the triggering epitopes of these cereals. Based on the sandwich ELISA results, the 20 ppm PWG gliadin standard (gluten free limit ≤20ppm) was involved to the indirect serological ELISA and all of the four einkorn genotypes were given lower absorbance, immunological signal than the 20 ppm PWG gliadin.

In the third year the 2D-GE and 2-D immunoblot analyses finished. Based on the results of the immunological investigations, from the 2D GE gels potential toxic and allergenic protein spots were picked. The results of the nanoLC-MS/MS workflow were very interesting.

In case of the four selected einkorn genotypes the gel digestions of the total protein extracts had contained in large amount saline soluble proteins instead typical prolamins.

These storage proteins mainly gave only the 10-20% of the total storage protein fractions, but recently proved the fact that some of them responsible for the development of the symptoms of

wheat related disorders. Attend of this knowledge it is not surprising, that based on the indirect serological ELISA assays the saline soluble fractions gave higher immune reactive response than the typical prolamins. Most of the scores the identified proteins were enzymes, stress related proteins like beta-amylase, serpin, serine protease inhibitor, and heat shock proteins, supplemented with the members of cupin superfamily, like globulins and vicilins.

From the prolamin scores the dominant components were LMW glutenins and gamma gliadins together with low amount of alpha gliadins and avenin-like storage proteins. In the original research plan did not contained the following analyses, we had much work with it, but the results strengthen the whole project, and gave very important details of the proteome and abiotic stress adaptation feature of the investigated einkorn genotypes.

In 2018 international cooperation of research teams from Hungary, Australia, Norway, Germany, and China has shed light on a promising development for patients suffering from celiac disease and wheat allergy. They examined proteins with a proven relationship to celiac disease, occupational asthma (baker's asthma) or wheat dependent exercise induced anaphylaxis (WDEIA), the results were published in the Science Advances. Within this framework of whet resequencing, all of the storage proteins were annotated by our group, including the wheat A genome, closest related to einkorn genome. We have developed the first complete representation of the proteins related the different forms of immune response in humans, which has helped us to accurately determine the genetic variability of these proteins and their environmental vulnerability. Along with mapping the location of these proteins on the wheat genome, the research team investigated how the environment affected the expression of proteins in developing grain, and resulting effect on human health.

In summary from the 200 investigated wild- and cultivated einkorn genotypes four proved to be low immunogen and low allergen protein level. Based on the commercially available G12 and R5 gluten test kits and serological ELISA all of the four flour were under the gluten free limit, but further preclinical (T lymphocyte proliferation test) and clinical (double blind placebo test) needed to prove that these einkorn's are safe for patients suffering in celiac and wheat allergy.

Partial results were published in the proceeding books of the Prolamin Working Group.

During the project there were some extra research steps, and shifting with the original research plan, due to the limitation of the human sera needed for the 2D-GE the immunoblot analyses and the five mouths long service period of the Crocodile ELISA workstation. Because of the shifting the final manuscript which summarizing the wet lab results will be published in this year in the Scientific Reports or in Frontiers in Plant Sciences.

During the project I had gained the International Youth Conference Competition in 2015 to travel to Valencia participate on Prolamin Working Group Meeting. I also had earned the Bolyai Research Scholarship of the Hungarian Academy of Sciences and the ÚNKP-18-4-BME-393 National Excellence Program of the Ministry of Human Capacities.

Wheat Initiative had elected me as a member of an expert group Improving wheat quality for processing and health, which was a great honour.