Understanding the impact of maternal diet and microbiota on celiac disease development in an at-risk population

-

Sophie Rivara

EPFL – May 2016

Student reviewer 1: Anne Billet

ABSTRACT

This study aims at better understanding celiac disease onset and characterization. Using a hybrid human observational-intervention trial, we want to assess celiac disease prevalence in children depending on their mothers' microbiota. For this, future mothers, either celiac or carrying main susceptibility variants for celiac disease, will be recruited. They will be subjected to various combinations of treatments: gluten free or normal diet, probiotics or placebo intake, breast or formula feeding of their children. Periodic blood and feces collections from the mothers and the children, as well as multiple survey fillings and follow-up meetings with clinicians and dieticians during at least five years, will allow the constitution of a huge database. First results drawn from this huge mass of information should reveal links between maternal diet, probiotics intake and microbiota. Furthermore, they should enable to correlate maternal microbiota (and hence nutritional behavior during pregnancy and breastfeeding) with onset of celiac disease in the children. Eventually, this study also aims at collecting as much information as possible relative to early celiac disease onset, its evolution and the implication of probiotics treatment, nutrition or any other relevant and quantifiable environmental factor. From this perspective, this project can be seen as a pilot study setting the foundations of a complete database about celiac disease. Such a database will allow detecting unknown correlations and will give access to a huge amount of detailed data valuable for future studies in the field of celiac disease.

INTRODUCTION

Overall description of celiac disease

Celiac disease (CD) is an inflammatory autoimmune disease affecting around 1% of the population with European ancestors[1]. It is an adverse reaction to dietary gluten that usually results in enteropathy, principally in the small intestine. Gluten is a family of cereal-derived proteins found in wheat (gliadin), barley (hordein) and rye (secalin)[2]. Individuals can be genetically predisposed to CD by carrying HLA-DQ2 and HLA-DQ8 haplotypes at the human leukocyte antigen class II locus, also called major histocompatibility complex II, MHC II[3]. In predisposed individuals who develop CD, gluten ingestion activates both innate and adaptive branches of the immune system, ultimately leading to the production of pro-inflammatory cytokines and proteases that harm the intestinal epithelium. This is responsible for villous atrophy and tight junctions' disruption. The increased epithelial permeability allows more gluten derived peptides to reach the lamina propria, where they further trigger an inflammatory response[4]. This vicious circle leads to chronic inflammation, eventually resulting in impaired nutrients absorption. The gut microbiota also appears to be affected in a vicious circle model: predisposed individuals have a specific microbiota that favors the onset of an adverse reaction to gluten, which in turn enhances inflammation and further modifies the microbiota toward more pathogenic or pro-inflammatory strains[1,5].

On the molecular level, untreated celiac patients exhibit a highly inflammatory intestinal phenotype, triggered both by gluten peptides ingestion and celiac associated microbiota: their epithelial and dendritic cells produce various pro-inflammatory cytokines that in turn activate intraepithelial cytotoxic lymphocytes[4]. All the active innate immune cells release more cytokines such as IL-1, IL-6, IL-15, IL-21, IFN-α and TNF-α, which trigger further inflammation and increase epithelial permeability[5]. Simultaneously, gluten peptides reaching the mucosa and lamina propria bind to tissue transglutaminases. There, they undergo deamidation, which gives them high affinity for HLA-DQ2/8 T cell specific receptor, leading to an adaptive immune response directed toward gluten peptides and tissue transglutaminases[1,2,6]. The activated T cells contribute to the inflammation by secreting IFN-γ that trigger matrix metalloproteases release from macrophages and epithelial cells, participating into mucosa remodeling, villous atrophy and intestinal epithelium increased permeability[6]. At the same time, dendritic and T cells allow B cells activation and differentiation into plasma cells and subsequent production of anti-deamidated gluten-derived peptides and anti-tissue transglutaminase secretory IgAs[3–5]. Next to those massive immune reactions, the microbiota dysbiosis observed in active celiac patients (not treated) shows the presence of more pro-inflammatory bacterial strains than in non-celiac or treated celiac patients. This dysbiosis is thought to contribute to the disruption of the equilibrium between regulatory and pro-inflammatory immune signals, participating to the disease phenotype both as a trigger and as a consequence. It has been proposed to be part of the trigger of innate auto-inflammatory reactions[1,5,6].

To summarize, although not all molecular mechanisms are known and understood in details, it appears clear that in celiac patients, gluten-derived peptides can cause a severe and chronic inflammation. This inflammation is characterized by a change in the microbiota as well as an autoimmune reaction (from innate and adaptive arms of the immune system), both being vicious circles leading to an increased self-sustained inflammation.

A broad range of symptoms can be found in celiac patients, ranging from no clear external sign to bloating, gastrointestinal pain, diarrhea, vomiting and exhaustion due to the lack of nutrients (particularly in children)[3]. The only treatment up to date is a strict life-long adherence to a gluten free

diet (GFD). However, patients who don't experience very strong symptoms or were diagnosed after 4 years old tend not to strictly follow the demanding GFD and regime compliance can be as low as 40%[2]. Even for patients willing to comply, problems arise from cross-contaminations and life quality can be impaired[6–8]. Furthermore, the microbiota is not restored to its normal ratio of potentially harmful/beneficial bacteria and at least a low level chronic inflammation persists in up to 50% of celiac patients[7] (the causes for these observations are unknown, but might involve contaminations or consequences of GFD on microbiota composition).

Therefore, research has been undergone both to better understand and prevent the disease onset and to develop a treatment that could alleviate or suppress the symptoms. On one hand, most recent intervention trials aim at modifying the microbiota toward a less inflammatory phenotype. Some researches evaluate the use of probiotics or helminth therapy to act on the microbiota[6–9]. However, even though those results appear successful at decreasing the inflammation in patients under GFD, nothing yet can replace the life-long compliance with this strict demanding regime. On the other hand, research aiming at preventing and understanding the disease onset has also been conducted. Because of clear microbiota implication, lots of groups have been addressing the effect of infant feeding. However, results are often controversial and many questions stay unanswered[10]. Hence, more research to understand CD triggers and how it could be prevented is needed.

Origins of celiac disease

Extensive studies aiming at determining the exact trigger of CD couldn't define a clear pathogenesis origin up to now. Genetic predispositions are involved, the clearest being that celiac individuals possess at least one MHC II genetic variant among HLA-DQ2 or HLA-DQ8. However, whereas up to 30-35% of the population is genetically predisposed, only 2-5% of the predisposed individuals develop CD, leading to an approximate 1% of the population being affected[4]. Hence an unknown trigger has to be involved in CD pathogenesis. Gluten ingestion is of course required, but the fact that onset can vary in time (not every celiac patient develops the disease at gluten introduction) tells us that other factors are involved[5]. In fact, although one doesn't know exactly how the various considered elements impact on CD onset, three different categories of factors have been proposed to play a role.

First, as mentioned, genetic predispositions are keys to the disease onset. Reviews report between 31%[1] and 54%[3] of CD heritability attributable to defined genetic variants. The genetic regions associated with CD contain the HLA locus and 39[1] (or 43[3]) non-HLA related regions. Those have been associated with various immune functions, but also bacterial sensing and colonization. Moreover, 81% of the genetic susceptibility regions are non-coding portions of DNA[1]. These intergenic regions are expression regulation loci affecting transcription levels of some genes[1,3]. Together, those studies indicate an important role of genetic and epigenetic factors in CD onset.

The second factor, the implication of the microbiota, might be the most investigated at the moment. Comparative analysis, intervention studies and literature reviews have been extensively conducted to interrogate what affects the microbiota and how the microbiota in turn affects CD onset and symptoms[1,2,4–9,11,12]. Although not all studies yielded the same findings, there is an almost general consensus to link active CD with dysbiosis in untreated patients and with persistent unbalanced intestinal microbiota in individuals treated by a GFD. Furthermore, some authors suggest the microbiota as the integrative point of other factors, proposing that environmental and genetic influences all affect the microbiome[1,4,5]. In turn, this microbiome would contribute to the balance between regulatory and inflammatory host immune reaction, explaining why a dysbiosis could lead

to dramatic effects in host immune regulation and inflammatory disorders such as CD. Notably, two recent studies aiming at characterizing the celiac associated dysbiosis concluded with drastically opposite answers that intestinal flora of active patients was[11] or not[12] markedly modified compared to healthy controls. These contradicting results and the lack of clear and consistent celiac-associated microbiota characterization leave room for subsequent studies in this area.

Finally, the last factors potentially impacting on CD onset are environmental ones. This category includes a broad range of putative factors such as birth delivery mode (caesarean or through vagina), breastfeeding or infant formula (mothers eating gluten could transmit gluten-derived peptides very early to their child through their milk[13], which also differs between CD and healthy mothers[14]), infections and/or antibiotics intake, timing of dietary gluten introduction… Here again, lots of studies have been conducted to address many questions. A particular focus has been placed in the elaboration of recommendations regarding breastfeeding and time of gluten introduction, stipulating that children should be breastfed and gluten should be introduced progressively between 4 and 6 months while still breastfeeding to minimize risk of CD development[15,16]. However recent studies showed that both breastfeeding and timing of gluten introduction had actually no influence on CD onset and prevalence[10,13]. A review based on meta-analysis of data from 21 publications[10] addressed the impact on CD of breast feeding and other environmental factors. The most striking conclusion is the discrepancy between the results gathered from the various publications selected. This lack of consistency brings forward the need for structured and well established standards to assess environmental factors in a less biased and hopefully more reproducible way.

In conclusion, while lots of studies are trying to unravel the mechanisms by which CD starts, no real consensus emerges and a lot remains not understood yet. What can be said so far is that CD seems to be caused by multiple intermingled factors including genetics and epigenetics, microbiota (gut colonization and later evolution) and environmental factors, especially during early childhood, when the immune system has still to mature. With the exception of genetic predispositions, factors favoring CD onset are controversial and deserve further research. Today, no satisfying treatment is available against CD and we lack a good understanding of the triggers. For celiac patients or non-celiac people bearing predispositions, the inability to “do the right thing” to prevent transmission to children is a real issue that needs to be addressed.

PROPOSED STUDY

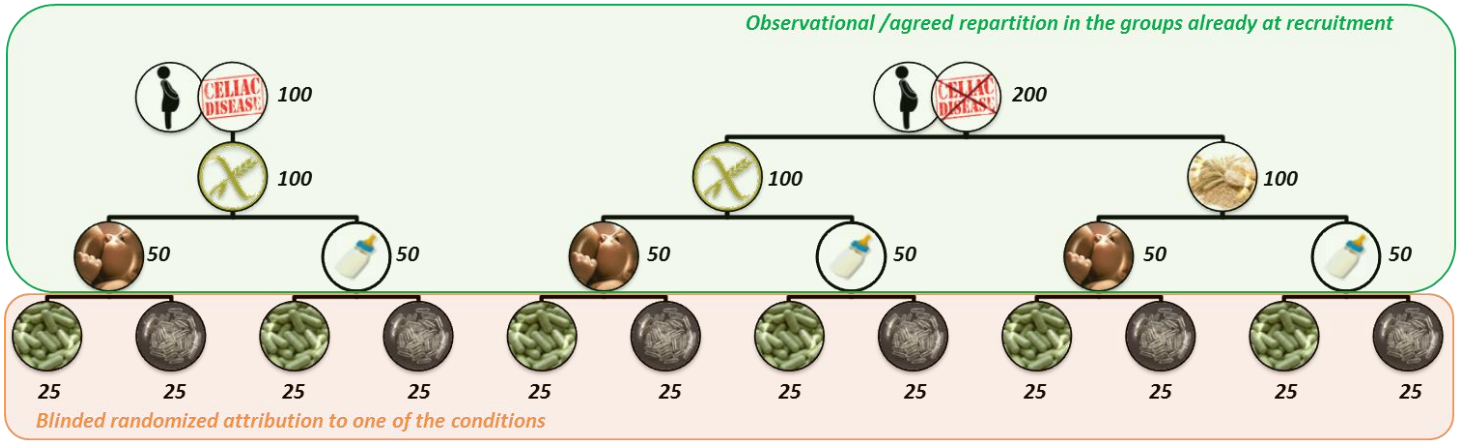
Aim

We suggest a study that would directly interrogate heredity of CD in order to see whether easy early interventions could prevent the disease onset. More specifically, we want to assess whether a GFD, the intake of probiotics by the mother during pregnancy and breastfeeding have an influence on CD development in an at-risk population. In addition, we want to structure our study so that we can assess many parameters and collect as much data as possible, in order to generate a database that could be of further interest in the field of CD research.

Experimental design and rationale

To pursue this aim, we want to design a human study combining observation and blinded randomized intervention in future mothers predisposed to CD, being themselves celiac or not. The study will follow the mother and her child if the child carries the main HLA genetic predisposition. At the moment, the most promising advances regarding treatment and understanding of CD seem to be related to the microbiota. However, no previous study addressed in a systematic way the effect of

the maternal microbiota. This is what our study will do, observing in two steps the effect of GFD and probiotics intake first on the mother microbiota and then how it translates into the child, depending on other environmental factors such as infant feeding. The final binary outcome of the study will be whether or not the children develop CD until the age of 5. When applicable, update of the results to include later CD onsets should be accomplished. Figure 1 depicts in details how groups will be constituted for the study.



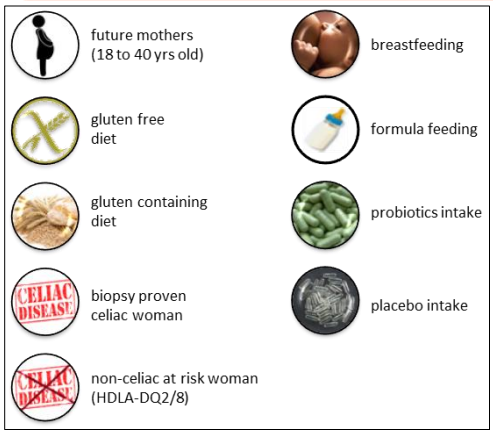


Figure 1. Schematic view of the experimental design. Maternal groups structure for the study, allowing for assessment of CD, GFD and mode of infant feeding. The last separation, based on probiotics intake, in contrast with the first criteria used to distribute the women among the groups, is a blinded randomized attribution to either probiotics or placebo intake groups. The group attribution for the first criteria is done at recruitment, upon discussion with the mother and her future plans. To ensure a reasonable number of celiac children in the study, the mother have to be celiac or at-risk and the future child has to have at least one celiac relative (first or second degree, either in the maternal or paternal family). The final binary outcome (i.e. development or not of CD) can be assessed in the children and results analyzed grouping the children according to the treatment followed by their mother. During data analysis, further separation of children is possible based on either their mother HLA-DQ subtypes (2 or 8) or their own. If other more predictive genetic predispositions are found, they can of course be used too. A supplementary output, the microbiota composition based on feces collection, is available for both the mothers (at various stages during pregnancy and breastfeeding) and the children (at many time points from birth to 5th birthday.
Numbers indicate approximatively how many women will be recruited per group.

Even if the first step fails to show any significant differences (i.e. if probiotics and/or GFD don't impact on maternal microbiota), the second step of the study, following the children, will be worth it. Indeed, the multiple data collected will allow for in-depth result analysis and leave the opportunity for unforeseen criteria to become determining. In fact, our goal is to set standards allowing our data to be used by future studies. For this reason, we collect as much information as possible on the participants while using methods that are as non-invasive as possible. Continuation of the study at a larger scale could be possible if any of the observed criteria show trends or significant differences that are of therapeutic or scientific interest. Indeed, the data collected would theoretically allow us to separate participants in up to 24 groups depending on the interrogated criteria (groups depicted on Fig. 1 can further be subdivided according to mother HLA loci). Hence, this study is designed as a pilot study with preliminary results allowing re-focusing on points of interest and future expansion into large scale studies or scientific collaborations.

Namely, in addition to GFD or not and intake of probiotics or not, participants will be subdivided according to the following binary criteria: CD of the mother, type of genetic predisposition of the mother (HLA-DQ2 or 8, other celiac-associated genetic loci could also be interrogated since the

information will be collected), formula or breastfeeding, type of genetic predisposition of the child. Out of all those criteria, most will be observed (celiac mother, genetic predispositions, compliance to a GFD, breastfeeding or infant formula), while effect of probiotics intake will be assessed in a blinded randomized way. Moreover, other non-binary data collected will include mothers and children genetic predispositions (through targeted sequencing at the known celiac-associated loci or complete sequencing depending on participant's consent) and multiple time-points feces samples allowing for microbiota characterization.

METHODS

First of all, the whole study workflow will have to be reviewed and acknowledged by the Swiss Research Ethics Committee and the clinical trial protocol needs to receive Swissmedic's approval. Once this is done, recruitment will start planed on a 3 years basis. Enrolled participants (see criteria below) will then go through the various steps described below (see also table 1).

Participants

For our study, we will recruit participants using general advertisement (in hospitals and universities or any public spaces where eligible women are susceptible to be) as well as targeted announcements (through clinicians and celiac patients associations). Candidates are celiac or at-risk for CD (presence of at least one of the HLA-DQ2 or HLA-DQ8 variants will be assessed on blood samples at enrolment) future mothers (planning to have children or pregnant from maximum 3 months at time of enrolment). Their child to come has to have a first or second degree celiac relative. They consent to enroll their future child into the study. A meeting with a clinician and a researcher will be organized, in which the study will be detailed. We will make sure to obtain informed consent from each participant in accordance to the law and ethics. Capacity of the future participant to comply with the procedure and contribute until the end of the study will also be assessed.

At enrolment, the participant has to be attributed to a study group. Discussions with an investigator will establish whether the mother plans to breastfeed her child or not and whether she wants to follow a GFD or not (for non-celiac women that have this choice). Finally, a blinded random binary decision will attribute the participant to the group consuming either probiotics or placebo. First steps of the study will start during pregnancy. At birth, analyses based on placental material will allow to determine the child's HLA variants. If the child doesn't carry one of HLA-DQ2 or 8, he'll be excluded from further analysis (however, data on the evolution of the maternal microbiota could still be used). An additional requirement is natural delivery at birth. If medical reasons force to use cesarean, the child will still be monitored but might be excluded from statistical analysis if he/she appears to be an outlier.

In addition, the investigator can remove a participant from the study cohort if one of the following exclusion criteria is met: highly premature birth (<32 weeks), development of any immune disease which could bias the results, non-compliance to the required diets/practices.

Experimental steps and data collection

The study is built as follows (see table 1 for a recapitulation of the various steps and the corresponding data collected):

At recruitment, candidates will have to fill an informed consent document in which decision about type of sequencing and results to be communicated are clearly established. With 10 mL blood samples, HLA locus will be sequenced to check for predispositions to CD. Retained participants will be

attributed to a group as described above. Furthermore, their blood will be used for further genomic analysis (depending on participant's consent, only targeted sequencing for regions previously associated with CD or whole genome sequencing are done) and screened for anti-gliadin and anti-transglutaminase antibodies (characteristics of an immune reaction against gluten in celiac patients) and for various pro-inflammatory cytokines. In addition, participants will learn to collect feces. Their feces will be collected (on site or in the upcoming days) to allow for further microbiota analysis (by 16S rRNA sequencing) as well as sIgA detection and identification. Finally, participants will meet with a dietician. Eating habits and average gluten consumption will be evaluated by discussion and survey filling by all groups, whereas advices and control of real gluten avoidance are defined for women in GFD groups (this diet must be followed for at least 6 months at the beginning of the pregnancy).

At 13 weeks of pregnancy, which corresponds to the beginning of the second trimester of pregnancy, participants start the probiotics or placebo treatment by daily pills intake before breakfast (start of treatment is defined to allow effect on intestinal flora before birth and hence to potentially impact on primary GI tract colonization in child with natural delivery). Exact probiotics will be chosen by the experimenters among the ones available on the market that seem to yield the most promising results and contain at least two strains of *Bifidobacterium*, a bacterial family found in reduced amounts in CD patients. Placebo will be similar pills containing no bacterial strains. This treatment is taken daily by the mother until the 8^{th} month of breastfeeding (or until birth of the child when no breastfeeding, according to the group attribution).

Breastfeeding or formula feeding (same formula provided to all not breastfeeding parents) is strictly regulated and follows the same protocol to avoid bias in the comparison. Both are accomplished at least until the child reaches 8 months. In addition, solid food introduction has to be progressive and not before 6 months, while gluten introduction, starting with small amounts, will begin at 7 months, so that it is not the primary solid food introduced and it still occurs while maternal/formula milk is consumed. Details on introductions will be monitored precisely in surveys.

Dietetic surveys, blood sample uptake and feces collection, for both mother and child, will be regularly done (see exact planning in table 1). Note that the high frequency of child feces collection is planned to ensure a tight monitoring of the intestinal microbiota formation and early evolution upon changes in diet. In addition, follow-up and screens for CD will be undergone at gluten introduction and then every 4 months up to 24 months of age, then every 6 months up to 5 years. Follow-up will then be encouraged and offered every year up to 16 years included, both to encourage responsible health monitoring and favor detection of late-onset CD, also allowing more long-term data to be collected. If CD is suspected, adequate diagnostic techniques will be used and, if biopsy confirms CD, a GFD will be started while follow-up of the children continues.

Finally, unrelated diseases, infections, treatments (with antibiotics or not), vaccines or other interventions that could potentially affect the immune system or GI integrity should be announced by the patient and monitored. If upon preliminary results, some criteria appear to be missing or biased and can be corrected for, adjustments can be made (e.g. add a feces or blood collection time-point between two measures that drastically differ to monitor smooth evolution…).

Note that most medical appointments can be grouped, so that patients don't have them too frequently (maternal and child meetings are synchronized and each appointments cumulates blood sampling, general monitoring by clinicians and meeting with the dietician).

Table 1. Recapitulation of the study. Different experiments are recapitulated. In addition, timing and frequency of information assessment and type of samples or data collected are indicated. Information missing in the table is that breast or formula feeding, as described in the main text, has to last for at least 8 months while solid food introduction starts at 6 months and progressive gluten introduction at 7 months; also, probiotics intake starts during the second trimester of pregnancy (week 13) and lasts until the 8th month of breastfeeding (if applicable, until child's delivery otherwise). (WA = when applicable, depending on group attribution and conditions)

Experimental step		Timing and frequency	Data/sample collected	Extracted information
In the mother	Blood sampling	At enrolment, at weeks 13 and 29 of pregnancy, at child's birth, 1 month post-partum, 4 months post-partum, then every 4 months until the end of breastfeeding (WA)	10 mL of blood	Genomic information (first sampling targeting especially celiac associated loci); anti-gliadin and anti-transglutaminase antibodies levels; anti-inflammatory cytokines (TNF-α…) levels; hematocrit; abundance of various metabolites
	Feces collection	Before pregnancy (WA), at weeks 13 and 29 of pregnancy, at child's birth, 1 week post-partum, 1 month post-partum, then monthly until the end of breastfeeding (WA)	Representative sample of daily feces, stored frozen in an appropriate plastic bags	Microbiota identification and microbiome sequencing (by 16S rRNA sequencing); monitoring of microbiota evolution; screen for and identification of sIgAs
	Diet monitoring	Weekly during pregnancy and until the end of breastfeeding (WA)	Surveys	Maternal diet (general food habits…)
		During follow-up meetings (concomitant with blood sampling) until the end of breastfeeding (WA)	Meetings with a dietician	Maternal diet details (amount of gluten and complement on general habits); good compliance to the GFD (WA)
	General monitoring	Together with other follow-up meetings (concomitants with blood samplings) until the end of breastfeeding (WA)	Meetings with a clinician	Assessment of potential perturbing factors (antibiotics intake, infections, travels…); monitoring of CD (WA)
In the child	Blood sampling	At birth, at months 1, 4 and then every 4 months until 24 months, every 6 months until 5 years and every year until 16 (WA)	Blood samples, 10 mL from placenta at birth, 5 mL per sampling until 24 months, 10 mL later on	Genomic information (first sampling targeting especially celiac associated loci); anti-gliadin and anti-transglutaminase antibodies levels; anti-inflammatory cytokines (TNF-α…) levels; hematocrit; abundance of various metabolites
	Feces collection	Daily during the first 10 days of life, weekly until 2 months, monthly until 12 months and every 3 months until 5 years	Representative sample of daily feces, stored frozen in an appropriate plastic bags	Gut colonization long-time monitoring; microbiota and microbiome characterization and monitoring (by 16S rRNA sequencing); screen for and identification of sIgAs
	Diet monitoring	Weekly until 12 months, monthly until 24 months and then every 3 months until 5 years	Surveys filled in by the parents	Information about food given to the child
		During follow-up meetings (concomitant with blood sampling) until the end of breastfeeding (WA)	Meetings with a dietician (and a parent)	Details of the food given to the child (including solid food and gluten introductions, caloric intake, type of food, gluten amounts…)
	General monitoring	Together with blood sampling	Meetings with a clinician (and a parent)	Assessment of potential perturbing factors (antibiotics intake, infections, travels…); CD monitoring (screening, follow-up if diagnosis…)

Compensations and help

Participants (mother and children) are followed for free by clinicians of the team (for what concerns CD and the study) and can obtain free meetings with a dietician upon request (the subjects to discuss have to be linked with the study). They are also taught how to collect and store feces by themselves (experimenters arrange with them to come at given intervals to collect samples). The experimental team is always available for questions and participants are encouraged to signal any element/event that they think could be of interest.

On the material point of view, the participants are provided the following, according to the needs of the study and their group attribution: material for feces collection and a freezer (for feces samples preservation and storage); a unique contribution of 1000 CHF for the supplementary costs of a GFD during pregnancy and breast feeding; probiotics or placebo; infant formula (provided by the experimenters, but not paid for, since this cost would have been there for the family anyway). An additional 3000 CHF retribution is given to mothers (1000 CHF for study completion up to child's birth and 2000 CHF upon complete study completion at child's 5th birthday).

EXPECTED OUTCOMES

As mentioned, this work can be considered a full study as well as a pilot study setting the basis and standards for a huge database constitution. While we will focus here on the direct readouts, one should not forget the global positive impact that all the collected data constitute for the field of CD related research. Also worth noting is the fact that, despite all the available information, analysts will need to focus and make choices of what they consider. Indeed, the variability inherent to every human being will generate some noise and hence, as for any human-based study, statistically significant correlations will have to be carefully established in the huge amount of collected data. We will here look more in-depth into outcomes that were primarily interrogated by our study and drove its conception.

The first readout will be the influence of simultaneous probiotics intake and pregnancy on celiac or non-celiac predisposed women. Data collected later on during the experiment should provide additional information about the effect of probiotics intake independently of pregnancy. While all kinds of comparisons could be foreseen, the first aim of this study will be to compare microbiota of women following a GFD or not during pregnancy, with or without probiotics intake. Hopefully, probiotics in celiac women will contribute to a lower basal level of pro-inflammatory cytokines and a higher proportion of potentially beneficial bacteria in the intestinal microbiota. Moreover, we expect a difference in the microbiota of GFD or non-GFD patients and it will be interesting to see which further changes to the microbiota appear when CD is present in addition to GFD.

The second direct readout will be to link mother microbiota with CD prevalence and onset in children. For this, infants will be grouped according to clusters in their mother's microbiota composition and significant prevalence of CD will be assessed. Supposing that the results show clusters for three different types of maternal microbiota at birth (associated with women who ate gluten, who followed a GFD with placebo and who followed a GFD with probiotics), we could observe how determining these criteria are for CD onset by comparing disease prevalence in children from those groups. Implication of breastfeeding could also be assessed, either in itself (separating the cohort in breast versus formula fed children) or in combination with maternal microbiota differences. Breastfeeding might yield an important effect, especially because milk composition should vary in mothers (impact of CD in itself or of non-GFD, gluten derivatives being then present in the milk). Another binary approach would be to separate the cohort with the different factors two by two (GFD or not, celiac or not, probiotics or not…) and see which of the criteria yields a significant difference in the prevalence of CD in the corresponding children. Even another output to look at, more complex (not binary), would be to try to correlate in details the maternal treatments with the intestinal colonization and microbiota evolution of the child over time, considering or not genetic background or other collected environmental factors (childhood infections, diet…).

No matter which output approach yields the better results, what is clear is that many questions could be addressed by this study and that using proper statistical analysis, lots of information could be retrieved and correlations established. With all the available factors and defined parameters to separate the cohort, many discrepancies could certainly be correlated with a given criteria. Indeed, one of the primary objectives of this study is to avoid the lack of coherence and contradictory results that have been so frequent up to now in the field by having enough parameters to detect and suppress potential bias.

Whatever they show, our results will be crucial for understanding and prevention of CD. Hopefully, they should allow for the establishment of harmonized guidelines. The discovery of "good practices"

to reduce celiac disease onset risk would eventually give the opportunity to predisposed or celiac people to try to "do the right thing" to prevent transmission of the disease to their offspring. Furthermore, the microbiota part could potentially contribute to the development of strategies to offer a better quality of life to celiac people.

PERSPECTIVE

Many uplifting perspectives would emerge with the potential realization of this study.

First of all, conclusions of this study will be central for a better orientation and focusing of research in the field of CD. They should contribute to unravel some of the mechanisms underlying disease onset. Of note, what could be learned about CD could probably help understanding and be applied to other inflammatory autoimmune diseases that are less well characterized, in particular in respect of what characteristics of the maternal microbiota are transmitted to the child and under which conditions.

Moreover, since neither completely new compound nor any drug is involved, discovery of some correlations (e.g. reduction of CD incidence in probiotics treated mothers) could quickly lead to revision of Health Organizations' recommendations, allowing for a quick important and efficient positive impact on CD prevention.

In addition, if probiotics show positive effects on the mother that don't seem to be transmitted to the child, a step forward would be to conduct an intervention trial in which some carefully selected probiotics would be mixed with maternal or formula milk or introduced right at the beginning of solid food consumption. Those interventions could consist in very interesting options, but would have the major drawback of being conducted on newborns or very young infants, with potential risks so high that lots of preliminary robust animal studies and modelling would be mandatory. Nonetheless, the present study is a first step toward better understanding of early celiac disease onset and microbiota constitution, which is the prerequisite of any intervention study aiming at impacting very young infants' microbiota. In this optic, it is an important step forward on the long road that could lead us to the development of a treatment or protocol allowing not to attenuate CD symptoms but directly to prevent its onset.

Eventually, as abundantly mentioned, the constitution of a structured, standardized and well documented database about CD would be a success in itself because it could positively impact all related research fields and help for many other studies. Hence, if promising results arise from this pilot study, further data collection and research, with focus on the most promising criteria, will be possible and accelerate tremendously CD related research.

TECHNICAL ASPECTS

This study will extend over 7 years (see figure 2) and be conducted by an interdisciplinary team as many different abilities are required. The principal investigator will run the team, attributing everyone's specific tasks and monitoring the global study progress. Since he will work in parallel on other projects, the part of his salary linked to this study will be in average 30%. An administrative assistant at 100% will manage all non-scientific tasks and scheduling. Recruitment meeting with the patient (including explanation of the study, signature of the informed consent form, first blood sampling, teaching of feces collection and material distribution) will require the presence of a gastroenterologist, a member of the research lab team, a nurse and a dietician (each for about 1 hour). For the hospital team, the most critical period will last 2 years (recruitment and, concomitant with last participants' recruitment, births of the first enrolled children), after which the work load will

decrease progressively. For 300 participants smoothly distributed over the recruitment period, a 20% dedication to the study during the first 2 years, 10% during years 2-4 and 5% until year 7, should be sufficient for the clinician, nurse and dietician. As soon as samples are collected, a PhD student can process them in the lab to extract the required data. Two life sciences PhD students will be employed for 5 years (one starting at the beginning with a focus on the maternal microbiota change, the other starting one year later with a focus on children microbiota monitoring). They will be assisted by master student interns (one at a time) and a lab technician (100% dedicated to this study). After 3 years, when data collection will have included all births, a third PhD student, doing bioinformatics, will start a 4-years project to analyze the collected genomes and microbiomes and finish with new samples processing. Furthermore, one data analyst at 50% (from year 2) will help with the huge amount of data and one computer scientist at 25% (from year 3) will build and manage the database. Finally a sample collector will be in charge of collecting the feces samples from the participants' freezer at strategic time points (about 300 half days of work over the 7 years).

Other costs include participants' compensations, material or products supplied to participants, lab consumables, use of platforms needed for the analysis, fees for analysis software and virtual storage space for the database, fees linked with the hospital and later children's follow-up. Of note, some of those costs will be effective also for candidates finally not enrolled in the study. Put all together, these costs reach an estimated total amount of 5'493'650 CHF on 7 years, so a bit less than 785'000 CHF every year (see details in table 2).

Table 2. Recapitulation and details of the total costs of the study. Salaries or various fees were estimated and calculated based on internet searches[17–21]. It is assumed that only 200 participants (mothers and children) will be fully sequenced (only targeted sequencing will be done for others). Abbreviations: C = candidates (not retained for the study); M = mothers; PP = pair of participants (mother + child); seq. = sequencing; SP = single participant (either mothers or children, separately).

Object/person		Monthly 100% cost / cost per use/sample (CHF)	# people – % work – # months/uses	**Total cost (CHF)**
Salaries	Principal investigator	12'000	1 person at 30% for 84 months	302'400
	Gastroenterologist	11'000	1 person at 20% for 24 months, 10% for 24 months and 5% for 36 months	99'000
	Nurse	6'000	1 person at 20% for 24 months, 10% for 24 months and 5% for 36 months	54'000
	Dietician	6'000	1 person at 20% for 24 months, 10% for 24 months and 5% for 36 months	54'000
	Lab technician	5'700	1 person at 100% for 84 months	478'800
	PhD students	4'300	2 people at 100% for 60 months and 1 person at 100% for 48 months	722'400
	Master students	1'000	12 people at 100% for 6 months each	72'000
	Administrative assistant	6'500	1 person at 100% during 84 months	546'000
	Data analyst	6'800	1 person at 50% for 72 months	244'800
	Computer scientist	6'800	1 person at 25% for 60 months	102'000
	Sample collector	120 for a half day	1 person 300 times	36'000
Facilities	Infrastructure (hospital)	25% * tot. hosp. team	-	51'750
	Whole genome seq.	1'500	200 SP	300'000
	Genomic targeted seq.	30	400 SP + 600 C	30'000
	Microbiota 16S rRNA seq.	30	49 samples * 150 PP + 70 samples * 150 PP	535'500
Material or other	Freezers	250	300 PP	75'000
	Probiotics and placebo	50	300 M * 6 months + 150 M * 8 months	150'000
	Compensation for GFD	1'000	200 M with GFD	200'000
	Participants' retribution	3'000	300 PP	900'000
	Software / server for database	2'000	60 months	120'000
	Lab consumables	5'000	84 months	420'000
Total				**5'493'650**

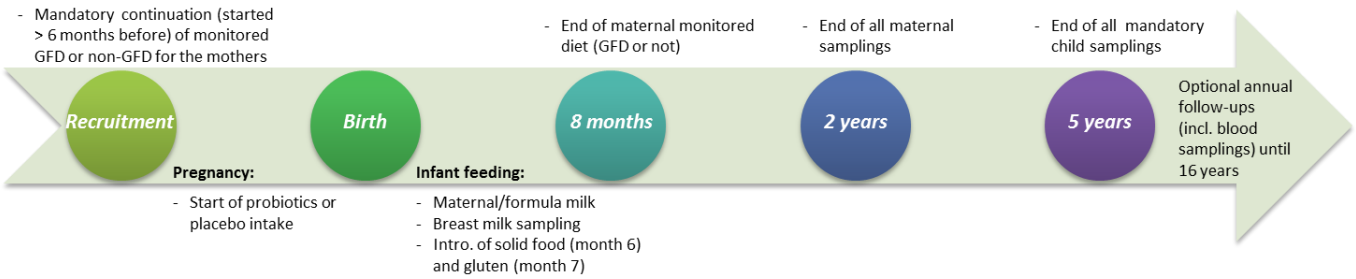


Figure 2. Timeline for the study. This scheme represents the several study steps and duration for one participant (mother and child). Because recruitment period should extend for about 15 months, the total duration of the study will be around seven years. Only main time points and study milestones are depicted on the timeline for readability reasons. However, frequent samplings for both mothers and children occur, as precisely described in table 1 above (feces collection, blood sampling, surveys for diet, CD screening or monitoring during follow-ups when applicable…).

CONFLICTS OF INTEREST

The author being celiac, discovery of any kind of treatment or prevention method represents a personal benefit. However, this shouldn't affect the study design and outcome, agreed that only strict scientific rigor can lead to any kind of usable results. Therefore, the author declares no conflicts of interest or biases in this study proposal.

REFERENCES

Scientific literature

1. Cenit, M. C., Olivares, M., Codoñer-Franch, P. & Sanz, Y. Intestinal Microbiota and Celiac Disease: Cause, Consequence or Co-Evolution? *Nutrients* **7,** 6900–6923 (2015).
2. de Sousa Moraes, L. F., Grzeskowiak, L. M., de Sales Teixeira, T. F. & Gouveia Peluzio, M. do C. Intestinal Microbiota and Probiotics in Celiac Disease. *Clin. Microbiol. Rev.* **27,** 482–489 (2014).
3. Turner, G. D., Dunne, M. R. & Ryan, A. W. Celiac Disease: Background and Historical Context. *Methods Mol. Biol. Clifton NJ* **1326,** 3–14 (2015).
4. Galipeau, H. J. & Verdu, E. F. Gut microbes and adverse food reactions: Focus on gluten related disorders. *Gut Microbes* **5,** 594–605 (2014).
5. Sanz, Y. Microbiome and Gluten. *Ann. Nutr. Metab.* **67 Suppl 2,** 28–41 (2015).
6. Marasco, G. *et al.* Gut Microbiota and Celiac Disease. *Dig. Dis. Sci.* (2016). doi:10.1007/s10620-015-4020-2
7. Klemenak, M., Dolinšek, J., Langerholc, T., Gioia, D. D. & Mičetić-Turk, D. Administration of Bifidobacterium breve Decreases the Production of TNF-α in Children with Celiac Disease. *Dig. Dis. Sci.* **60,** 3386–3392 (2015).
8. Olivares, M., Castillejo, G., Varea, V. & Sanz, Y. Double-blind, randomised, placebo-controlled intervention trial to evaluate the effects of Bifidobacterium longum CECT 7347 in children with newly diagnosed coeliac disease. *Br. J. Nutr.* **112,** 30–40 (2014).
9. Giacomin, P. *et al.* Experimental hookworm infection and escalating gluten challenges are associated with increased microbial richness in celiac subjects. *Sci. Rep.* **5,** 13797 (2015).
10. Szajewska, H. *et al.* Systematic review with meta-analysis: early infant feeding and coeliac disease--update 2015. *Aliment. Pharmacol. Ther.* **41,** 1038–1054 (2015).
11. D'Argenio, V. *et al.* Metagenomics Reveals Dysbiosis and a Potentially Pathogenic N. flavescens Strain in Duodenum of Adult Celiac Patients. *Am. J. Gastroenterol.* (2016). doi:10.1038/ajg.2016.95
12. Nistal, E. *et al.* Study of duodenal bacterial communities by 16s rrna gene analysis in adults with active celiac disease versus non celiac disease controls. *J. Appl. Microbiol.* (2016). doi:10.1111/jam.13111
13. Chmielewska, A., Pieścik-Lech, M., Szajewska, H. & Shamir, R. Primary Prevention of Celiac Disease: Environmental Factors with a Focus on Early Nutrition. *Ann. Nutr. Metab.* **67 Suppl 2,** 43–50 (2015).
14. Olivares, M. *et al.* Human milk composition differs in healthy mothers and mothers with celiac disease. *Eur. J. Nutr.* **54,** 119–128 (2015).
15. Guandalini, S. The influence of gluten: weaning recommendations for healthy children and children at risk for celiac disease. *Nestlé Nutr. Workshop Ser. Paediatr. Programme* **60,** 139–151; discussion 151–155 (2007).

16. Ludvigsson, J. F. & Fasano, A. Timing of introduction of gluten and celiac disease risk. *Ann. Nutr. Metab.* **60 Suppl 2,** 22–29 (2012).

Webpages consulted for costs estimation:

17. SCRIS ::::: Calculateur de salaire en ligne pour le canton de Vaud. Available at: http://cms.unige.ch/ses/lea/oue/projet/salaires/scris/. (Accessed: 6th May 2016)
18. Doctoral student's salary | PHD | EPFL. Available at: http://phd.epfl.ch/salary. (Accessed: 6th May 2016)
19. Pricing - Genomic Sequencing and Analysis Facility User Support Wiki - UT Austin Wikis. Available at: https://wikis.utexas.edu/display/GSAF/Pricing. (Accessed: 6th May 2016)
20. What Does It Cost to Run a Small College Lab? *Uncertain Principles* (2011). Available at: http://scienceblogs.com/principles/2011/08/16/what-does-it-cost-to-run-a-sma/. (Accessed: 6th May 2016)
21. Dedicated Server Hosting & Management | Rackspace. *Rackspace Hosting* Available at: https://www.rackspace.com/de-ch/dedicated-servers. (Accessed: 6th May 2016)

Sources of images used for figure 1

http://pictogram-free.com/45-act/172-act.html
http://bellyblog.ca/whats-all-the-fuss-about-gluten-celiac-disease-101/
https://commons.wikimedia.org/wiki/File:Gluten_free_SVG.svg
https://www.santenatureinnovation.com/devez-vous-arreter-le-gluten/
https://www.amazon.co.uk/Alchemists-Apothecary-Empty-Gelatine-Capsules/dp/B007L3BQ48
http://www.shrinelifesciences.com/dietary-supplements.htm
http://www.ohbebe.net/les-6-etapes-pour-preparer-un-biberon-comme-une-pro/
http://www.independent.co.uk/life-style/health-and-families/health-news/breastfeeding-boom-could-save-nhs-40m-a-year-8215599.html