REPORT SHORT TERM SCIENTIFIC MISSION
COST Action: FA1408

STSM title: **Application of next generation sequencing for whole genome analysis of different foodborne protozoa**

Reference : ECOST-STSM-FA1408-200616-079073

STSM dates: from 20-06-2016 to 22-07-2016

Location: **University of Warwick, Coventry, United Kingdom**

Host: Prof Mark Pallen, University of Warwick, m.pallen@warwick.ac.uk

COST STSM SCIENTIFIC REPORT

1. Aim and objectives of the STSM

Food-borne diseases are one of the most serious public health problems and one of the leading causes of illness and deaths[1-3]. Each year in the United States nearly forty-eight and a half million episodes occur associated to food[4]. Among them, *Toxoplasma gondii* causes the fourth highest number of hospitalizations and the second largest in deaths and human costs associated carry over 2,000 million dollars a year[5]. Similar results are exhibited by other countries as UK[6] and Australia[7].

We focus on *Toxoplasma gondii,* an obligate intracellular protozoa capable of infecting a wide variety of hosts in most edible tissues, and epidemiological studies has pointed out that consumption of cured meat as a risk factor for acquiring toxoplasmosis during pregnancy which is associated with neonatal abortion and sickness[8]. It is estimated that 30% of the world's population may be infected by this organism and a committee of experts of the FAO-WHO has established as the fourth priority among foodborne parasites[9].

The aim of this STSM was to evaluate the presence of *T. gondii* using metagenomics approach to identify genome data and associated risk factors particularly in slaughtered pigs. The main objective was to learn the methodology for the detection and characterization of *T. gondii* by means of pyrosequencing using a metagenomics approach and learn to bioinformatics analysis.

2. Description of the work undertaken

The work had been undertaken in the Microbiology and Infection Unit, at the Warwick Medical School, in the University of Warwick, Coventry CV4 7AL (Warwick, United Kingdom). University of Warwick.

DNA extraction was performed from pig tissue in a sterile area (Figure 1) following an extraction protocol based on lysis with CTAB and purification using QIAGEN columns. Sixteen samples were analyzed per batch and then we follow a modified protocol of Illumina library preparation. Two different kits were assayed, the TrySeq and the Nextera Illumina kits according to the manufacturer instructions with modifications in the general laboratory (Figure 2A). Briefly, it consists of breaking DNA into small fragments, a step for end repair, then clean up without size selection followed by adenylation of 3' ends with an A-tailing mix and ligate adapters. Fragments are cleaned up with Sample Purification Beads twice and next day the fragments must be enriched and cleaned up again. Libraries were validated by quantification of DNA amount using Qubit and quality assay using Bioanalyzer chip. The libraries were run in a MySeq device (Illumina) (Figure 2B).

A **B**

Figure 1. A. Entrance to the sterile area of the laboratories in the Life Science building of the Medical School of the University of Warwick. **B.** Dress to work in the specific area.

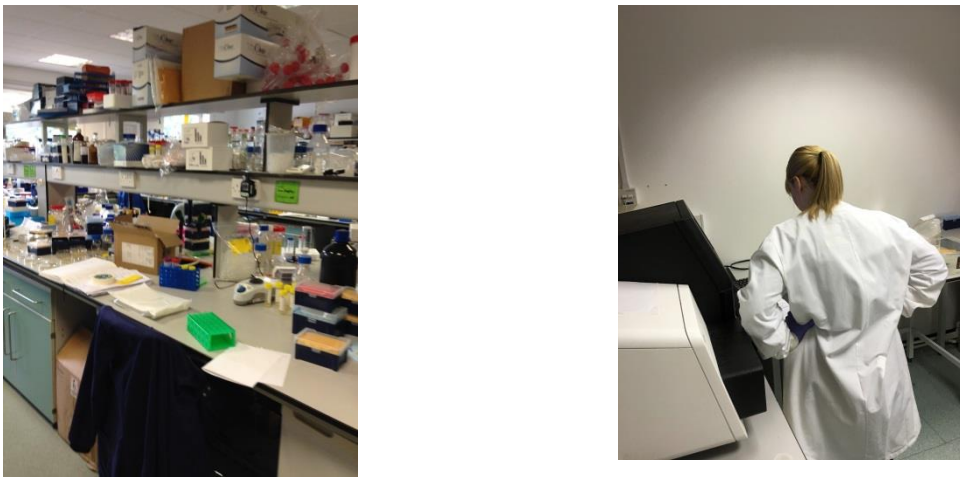

Figure 2. A. View of the general laboratory where samples and PCR were processed at certain steps and Illumina device **B.** (right) being launched with the assistance of Dr. Gemma Kay.

After a unique run that takes 2 days, 4.9 Gb of data were generated by run (Figure 3). Almost 80% of the clusters raised quality parameter control Q30 and 98% of the OTUs were identified. Another 3 runs were performed. Data were analyzed in a local server following bioinformatics pipelines already described by the group, particularly with the help of Dr. Oliver Smith. Several programs were used to analyze the data on a linux computer operating system (OS). I was familiarized with that OS learning different commands to call a program, create files, folders, count bases, create databases, etc. Firstly, we run BLAST over the reads that last a week, then we made first assignments using MEGAN, but better expertise is needed to data interpretation and analysis to avoid the platypus problem. I have also tried Nullarbor to look at most of the clinically relevant features and still bioinformatics analysis in being carried out. A strong collaboration has been stablished and my PhD student is now there in Warwick following the work done in June and July by myself.

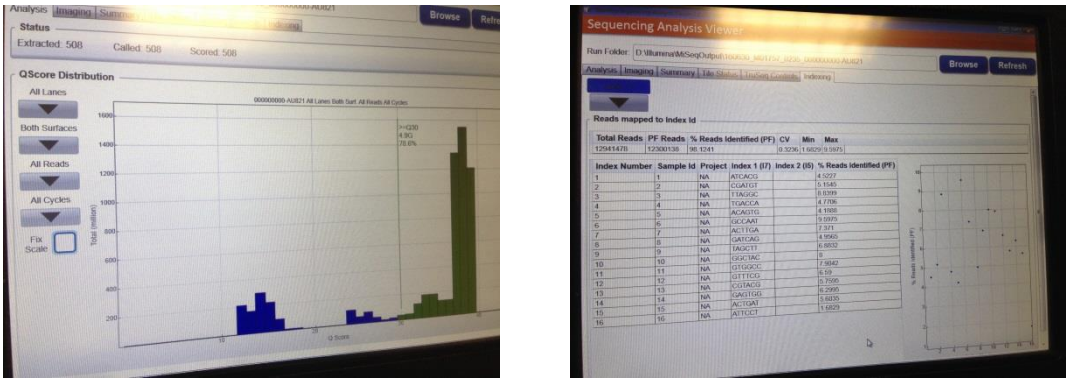

Figure 3. Data of quality of the results and mapped reads in the Illumina MySeq screen.

During the stay, Prof. Pallen who leads the CLIMB project (Cloud Infrastructure for Microbial Bioinformatics) in collaboration between Warwick, Birmingham, Cardiff and Swansea Universities, host the first meeting and training course. So, during 14th and 15th of July 2016, I had the opportunity to attend the meetings and a brief training course (Figure 4). I was also invited to the final dinner and the meeting taking in consideration not just the knowledge acquired but also the contacts performed, was really useful and attractive to me.

Figure 4. Reception of the participants in the CLIMB launch meeting in the hall of the Medical School building performed by the project manager Dr. Isabel Dodd. Right: participants attending the practical teaching in the informatics rooms.

3. Future collaboration possibilities with the host institution

This STSM has provided me new opportunities to study in collaboration with the host institution, University of Warwick, but also to develop good personal relationship with Prof. Mark Pallen. I also had the opportunity to contact other researchers thanks to this STSM as Tom Connor at the University of Cardiff or Nick Loman working at the University of Birmingham, and former collaborator of Prof. Pallen. During my stay, I have developed the skills necessary for doing metagenomics analysis in the lab and first steps of the further bioinformatics analysis. Due to the lack of time to repeat the doubtful results and also to do the long bioinformatics analysis, the results are still ongoing.

The data analysis is ongoing, indeed with the participation also of my PhD student, and after completion of this STSM, all the knowledge acquired in the host institution is being transferred in my home institution to the students and researches in the University of Burgos. Moreover, research findings are planned to be published in an international journal.

Acknowledgements

I would like to express my sincerely appreciation to Prof. Mark Pallen for this opportunity and his personal care during the STSM. And very special thanks to Dr. Gemma Kay and Dr. Oliver Smith who was training me in the laboratory and in front of the computer, respectively.

References

1. Wallace DJ, Van Gilder T, Shallow S, Fiorentino T, Segler SD, Smith KE, Shiferaw B, Etzel R, Garthright WE, Angulo FJ, and FoodNet Working Group (2000) Incidence of Foodborne Illnesses Reported by the Foodborne Diseases Active Surveillance Network (FoodNet)-1997. *J. Food Prot.* **63**: 807–809.
2. Anónimo (2010) Surveillance for Foodborne Disease Outbreaks: United States, 2007. *Morb. Mortal. Wkly. Rep.* **59**: 973-979.
3. Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson MA, Roy SL, Jones JL, and Griffin PM (2011a) Foodborne illness acquired in the United States--major pathogens. *Emerg. Infect. Dis.* **17**: 7-15.
4. Scallan E, Griffin PM, Angulo FJ, Tauxe RV, and Hoekstra RM (2011b) Foodborne illness acquired in the United States--unspecified agents. *Emerg. Infect. Dis.* **17**: 16-22.
5. Hoffmann S, Batz MB, Morris JG. (2012) Annual cost of illness and quality-adjusted life year losses in the United States due to 14 foodborne pathogens. *J Food Prot.* **75**: 1292-1302.
6. Adak GK, Meakins SM, Yip H, Lopman BA, and O'Brien SJ (2005) Disease risks from foods, England and Wales, 1996–2000. *Emerg. Infect. Dis.* **11**: 365–72.
7. Hall G, Kirk MD, Becker N, Gregory JE, Unicomb L, Millard G, et al. (2005)Estimating foodborne gastroenteritis, Australia. *Emerg. Infect. Dis.* **11**: 1257–64.
8. Iovu, A., Titilincu, A., Mircean, V., Blaga, R., Bejan, A., Cozma, V. 2008. Serosurvey of Toxoplasma gondii in sheep for human consumption in two slaughterhouses. Bulletin UASMV, Veterinary Medicine, 65(2), 40-43.
9. FAO/WHO (2014) Multicriteria-based ranking for risk management of food-borne parasites. Microbiological Risk Assessment Series No. 23. Rome. 302pp