Final publishable summary report

Executive summary

Chagas disease, caused by the protozoan *Trypanosoma cruzi,* is considered to be the most important parasitic disease in Latin America. The epidemiology of Chagas disease is complex, with distinct genetic lineages within the single species *T. cruzi.* The focus of this ambitious multidisciplinary ChagasEpiNet project, involving 15 research partners, was to apply high resolution technologies to elucidate the epidemiology of the genetic lineages of *T. cruzi*, for improved understanding and control of Chagas disease. Achievements include:

1. The development of techniques such as multilocus sequence typing (MLST) and high resolution microsatellite analysis (MLMT) for *T. cruzi* and their deployment in endemic regions have made widely available the capacity to identify *T. cruzi* genetic lineages, to unravel local transmission cycles and to plan local control strategies accordingly. Standardised PCR-RFLPs, MLST and MLMT methods have all been produced, with supplementary methods of data analysis, and technologies transferred, with high impact outputs. The detailed description within ChagasEpiNet of the ecologies and epidemiologies of the six distinct genetic lineages of *T. cruzi* (TcI –TcVI) has provided an essential and fundamental platform for all subsequent research on the epidemiology of Chagas disease.
2. The TcI genome has been sequenced and several unforeseen additional genome sequences also provided.
3. A special achievement has been the first development and deployment of peptide based lineage-specific serology.
4. Development of PCR oligochromatography proceeded as planned but it was not deployed to the field pending resolution of inconsistent results from different sources.
5. A large collection of clinical samples has been assembled, and pilot comparisons of genotyping and lineage-specific serology performed.
6. There has been an astonishing combined effort by South American and European partners to unravel in great detail the complex molecular epidemiology of *T. cruzi* infection, particularly in Bolivia, Brazil, Colombia, Ecuador and Venezuela, including ecological associations of the DTUs, domestic and sylvatic cycle interactions, congenital transmission and oral outbreaks, also revealing extensive inter-lineage mitochondrial introgression and intra-lineage genetic exchange.
7. Cell invasion, morphometry, pathogenesis, congenital transmission and drug susceptibilities of the DTUs have been compared.
8. Resolution and congruence of DTU identification and population genetics methods have been assessed, with phylogenetics analysis of relationships and evolution.
9. A greatly expanded international cryobank has been established.

ChagasEpiNet has on several objectives exceeded expectation, transforming research capacity and the understanding of Chagas disease. Aspects of the research warrant further funding, for example: continuity and expansion of partnerships on molecular epidemiology, further development of peptide based specific serology, and more sensitive, extensive analysis of infecting genotypes against clinical outcome.

Project context and objectives

Chagas disease is considered to be the most important parasitic disease in Latin America. The disease is caused by the protozoan *Trypanosoma cruzi*, which is transmitted primarily by bloodsucking triatomine bugs, not by their bite but by contamination of mucous membranes or abraded skin with the *T. cruzi* infected bug faeces and urine, deposited on the host as the bugs feed. Other methods of transmission occur, including transfusion of *T. cruzi* infected blood, transplantation of infected organs, congenital transmission, and oral transmission by contamination of food with infected triatomine faeces.

Control campaigns against triatomine bugs, by spraying infested houses, together with screening of blood and organ donors, have reduced the transmission of *T. cruzi* and the prevalence of infection in the human population. Nevertheless, it is still considered that 8 to 10 million people in Latin America carry *T. cruzi* infection. Mortality in the acute phase of infection is around 5 to 10%, with infants and young adults being the most vulnerable. With rare exceptions, once an individual is infected with *T. cruzi* they are infected to life, despite the immune response to infection. Approximately 30% of those who survive the acute phase of infection will progress to severe chronic Chagas disease. The organism multiplies in cells of a wide range of tissues, most notably in terms of pathogenesis, in muscle of the heart and alimentary tract, where both muscle cells and neurones are destroyed. Thus, in chronic Chagas disease there is cardiomyopathy with ECG abnormalities, particularly right bundle branch block (RBB) associated with sudden death or progressive congestive heart failure. A proportion of patients with chronic Chagas disease may develop chagasic megaoesophagus and or megacolon, requiring surgery. *T. cruzi* is also opportunistic, such that carriers of infection who become immunocompromised by HIV/AIDS may relapse into the acute phase. Both immunocompromised cases and congenital cases may be associated with meningoencephalitis. There is no vaccine against Chagas disease and no immunotherapy. There are only two drugs for chemotherapy, benznidazole and nifurtimox, which are not always curative, and both may have toxic side-effects, especially when administered to adults.

T. cruzi infection used to be considered to be almost entirely confined to Latin America. However, with increasing global migration from Latin America it is now estimated that 300,000 individuals carry *T. cruzi* infection in the USA and that tens of thousands of infections are present in several European countries. Cases of blood transfusion transmission and congenital transmission can arise far from Latin America. Thus the World Health Organisation (WHO) has launched a global awareness campaign to enhance recognition of Chagas disease outside the core endemic regions.

The epidemiology of Chagas disease is complex. In some endemic regions the triatomine bug vector species are confined to domestic and peridomestic sites, having arrived by dispersion or carriage from their original sylvatic locations. In other regions the same triatomine species may be present inside and outside the houses, with the threat of reinvasion from sylvatic habitats after spraying programmes. In the Amazon basin and the USA, although *T. cruzi* infection is present (enzootic) in a wide range of mammals the local bug species do not colonise houses and cases of vector borne Chagas disease are sporadic. Oral outbreaks occasionally occur in the Amazon region and elsewhere, particularly associated with consumption of triatomine contaminated plant juices.

T. cruzi used to be considered to be a relatively homogeneous entity, and it is still a single species. However, the application of biochemical methods, particularly multilocus enzyme electrophoresis (MLEE) had a remarkable impact on perceptions of the agent of Chagas disease. In a landmark study, in a village in Bahia State Brazil, MLEE demonstrated that in that focus domestic and sylvatic strains were radically distinct, more so than recognized species of the agents of leishmaniasis (*Leishmania*) species. Follow-up research with MLEE revealed that there are at least six distinct genetic groups or lineages of *T. cruzi.*

Molecular methods have advanced greatly in the last two decades with the development of higher resolution technologies based on the amplification and analysis of DNA. These technologies have profound potential impact on the understanding of the complex epidemiology of Chagas disease and of other infectious diseases, in the context of strategies for disease prevention and control.

The focus of this ambitious multidisciplinary project, involving 15 research partners, was therefore to elucidate the epidemiology of the genetic lineages of *T. cruzi*, for improved understanding and control of Chagas disease. The project united skills in genotyping, genomics, genetics and pathogenesis in Europe with considerable compatible skills in South America, and with key research in endemic areas that have distinct characteristics. The project was intended to be high impact in terms of both research progress and the fostering of collaborative networks.

Aim

Elucidate the epidemiology of the genetic lineages of *T. cruzi*, for improved understanding and prevention of Chagas disease.

Objectives

Technology development:

1. Develop further and apply genotyping technologies including PCR-RFLP typing, multilocus sequence typing (MLST) and multilocus microsatellite typing (MLMT) to the analysis of genetic populations of *T. cruzi*, making available improved standardised genotyping methods for routine use in endemic countries.

2. Sequence and annotate the unresolved genome of *T. cruzi* I

3. Guided by comparative genomics of *T. cruzi* I and *T. cruzi* VI (CL Brener strain), assess lineage-specific diagnosis with the trypomastigote small surface antigen (TSSA) and with other candidate antigens.

4. Develop a rapid PCR-oligochromatography test for detection of *T. cruzi* infection.

Molecular epidemiology:

5. Pilot studies of association between genetic lineage, clinical outcome, and prevalence of congenital infection

6. Assemble new isolates and biological clones of *T. cruzi* to map the silvatic vector, silvatic mammal, human and ecological associations of the *T. cruzi* lineages II, IV, V, VI.

7. Compare lineage specific pathogenesis and transmissibility of congenital infection in a mouse model, and compare lineage susceptibility to drugs *in vitro*.

Population genetics and phylogenetics:

8. Based on an expanded set of molecular data re-evaluate the population genetics and evolution of *T. cruzi* lineages.

International cryobank and database:

9. Establish in South America an accessible, expanded, international cryobank for *T. cruzi*

10. Establish a website and database for outputs of the project.

The project objectives encompassed the desirable characteristics prescribed by the relevant EC call, in that they included: genomics; effective, innovative relevance to disease, pathogenesis, drugs, interventions; an integrated multidisciplinarity, and capacity building, networking and training in endemic regions.

As this was a complex project with multiple objectives, for clarity the work performed since the beginning of the project and the main results achieved so far are summarised here sequentially for each of the 15 workpackages that comprised the project. The project ran for four years, with the fourth year being a no additional cost extension. Outputs of the research will emerge beyond the four-year life of the ChagasEpiNet project as data analyses and publications continue to be produced.

Main results and outputs

A meeting and a workshop planned for Bogota, Colombia in early 2012 was moved to Goiania, Brazil in September 2012, to combine a plenary scientific and financial meeting with three weeks of intensive research in the local laboratories.

Four PCRs were selected for DTU assignment and specificity to *T. cruzi* was confirmed. DTU assignment by PCR-RFLPs was highly congruent with MLST (> 99%). Standardised PCR-RFLP technology was transferred to the project partners and widely applied in endemic areas. Trials were also run on DNA extracts of blood/guanidine samples from chronic patients. A final MLST genotyping scheme has been optimised. One combination of 7 genetic loci discriminated all reference strains. A reduced scheme based on 4 gene fragments also displayed high bootstrap values for all DTUs. In parallel a mitochondrial MLST (mMLST) scheme with 10 targets has been developed and applied in endemic regions. Partner 10 developed MLSTest, a novel Windows based software for MLST data analysis in eukaryotes (http://www.ipe.edu.ar/). Advanced training in MLMT analysis has been provided. Publications included comparison of MLMT and other genotyping analyses, transmission dynamics in Colombia, definition of the origin of two recent oral outbreaks of disease in Venezuela, and the presence of recombination within sylvatic populations of *T. cruzi* in Bolivia.

The sequence of the Sylvio X10/1 genome was published and sequencing strategy improved to include Illumina data. Four additional strains, *T. cruzi marinkellei*, a bat specific

T. cruzi, a TcIV strain from Venezuela and two additional TcI strains have been sequenced. Comparative genomics has been used to identify polymorphic genes that might provide diagnostic tools. ELISAs using lineage-specific peptides were performed in Goiânia, Brazil with sera from Brazil, Ecuador, Venezuela, Argentina, and Bolivia. Collaboration on serology was established with the American Red cross. A promising pilot study was undertaken in Goiânia to explore biomarkers for distinction of cure and relapse. Cyokine assays suggested a role in presentation of chagasic cardiomyopathy, and levels were significantly different in TcI and mixed TcI/TcII. Further development of the PCR-oligchromatography test was shelved pending resolution of inconsistent specificity and sensitivity results with samples from different sources.

Goiânia was established as the site of the reference library for DNA analyses and serology, with samples from 720 individuals in six partner countries. A modified haemoculture procedure attained a tenfold increase in parasite isolation to >30%. As planned, pilot studies of genotype and lineage- specific serology have been performed.Intensive field research continued in Bolivia, Venezuela, Argentina, Brazil, Colombia and Ecuador. A combination of PCR-RFLPs, MLST and MLMT were applied to analysis of epidemiology, with several resultant publications.

Invasion and intracellular propagation assays have been completed for *T. cruzi* strains, and for *T. cruzi marinkellei*, in bat, opossum, human and monkey cell lines. All six lineages have been studied in IFNγR1-/- animals: differences exist between the *T. cruzi* DTUs at both the immune response and organ damage levels. Whatever the *T. cruzi* DTU, in a mouse model acute and chronic infections prejudice gestation outcome; congenital infection is rare and breast milk transmission unlikely. However, congenital transmission of TcV, most commonly associated with human congenital cases, was demonstrated in the mouse model. Fifteen *T. cruzi* strains representating the lineages displayed differential susceptibility to new drugs from *Streptomyces* sp. An algorithm summating drug effects at different concentrations, with cluster analysis, discriminates strains into at least four groups resembling the DTUs.

Population genetics and phylogenetic analyses have been performed on an unprecedented scale, using comparative MLST, mitochondrial MLST and MLMT, with addition of principal components analysis and dating of emergence events. The *T. cruzi* cryobank linked to the COLTRYP FIOCRUZ collection has been expanded further, with more than 100 sylvatic and human isolates (http://www.cria.org.br/~sidnei/fiocruz/coltryp/index). The established website (www.ki.se/chagasepinet) is linked to the cryobank catalogue (http://www.cria.org.br/~sidnei/fiocruz/coltryp/index) and the ChagasEpiNet sample database (https://creator.zoho.com/panstrongylus/chagasepinet-sample-database/).

Potential impact

The *socioeconomic* and *societal* impact of ChagasEpiNet is abundantly apparent from the nature of the research undertaken and the outputs summarised above. Thus, technology has been developed specifically in the context of improving understanding of the epidemiology and the control of Chagas disease. Research by all the partners in the endemic regions has been conducted to define transmission pathways and guide better strategies for disease control, for example in showing whether reinvasion is likely after spraying campaigns or monitoring and tracking emergence of new risks. The partners in the endemic regions are either an integral part of or closely affiliated to the public health systems. Thus outputs of this research are and will impel modification of vector control strategies. Awareness and surveillance for congenital Chagas disease have been promoted, for example in Venezuela and Ecuador. Impact has encompassed technological innovation, capacity, understanding,

awareness, process, performance, public health and policy. A very important impact has been the cementing of synergistic collaboration between the South American partners and the building of strong mutually supportive interactions between Europe and South America. New partnerships have arisen in South America as a result of the ChagasEpiNet initiative. A few examples of impact are reiterated below:

1. The detailed description within ChagasEpiNet of the ecologies and epidemiologies of the six distinct genetic lineages of *T. cruzi* (TcI –TcVI) has provided an essential and fundamental platform for all subsequent research on the epidemiology of Chagas disease.

2. The development of techniques such as multilocus sequence typing (MLST) and high resolution microsatellite analysis (MLMT) for *T. cruzi* and their deployment in endemic regions has made widely available the capacity to identify *T. cruzi* genetic lineages, to unravel local transmission cycles and to plan local control strategies accordingly.

3. In Ecuador population genetics analysis indicated rapid dispersal of domestic/peridomestic *T. cruzi* and genetic exchange among the domestic/peridomestic *T. cruzi* strains. Mapping of endemic localities that are predominantly separated from the sylvatic cycle has facilitated vector control.

4. In Venezuela TcI was confirmed as the primary agent of Chagas disease. TcIV is a secondary, putatively more benign disease agent. A relatively homogeneous clone of *T. cruzi* was shown to be widely dispersed. Nevertheless, sporadic invasion of sylvatic *Rhodnius prolixus* and *T. cruzi* strains demanded modified spraying programmes incorporating longer term surveillance. Oral outbreaks were shown to be due to TcI from contaminating sylvatic triatomines. Recommendations were made for the limitation of oral outbreaks.

5. In Colombia complex patterns of Tc1 co-infection were revealed with genetic exchange among strains, congenital transmission and oral outbreaks that involved sylvatic strains.

6. *T. cruzi* lineage-specific serology with synthetic peptides has been deployed in Goiania, Brazil to identify patients infected with TcII/TcV/TcVI. With David Leiby of the USA, lineage-specific serological screening of Latin American migrants has been successfully trialled to identify high risk blood donors. A feasibility study to screen Latin American migrants to UK has also been conducted.

Dissemination has primarily been by the generation of a large number of publications in scientific journals, and this dissemination is evident from the list of publications provided separately in this final report. Further publications are in preparation. Wider scientific dissemination has been by representation of ChagasEpiNet at many international conferences through oral presentations and multiple poster presentations. In addition, there has been discussion and interaction with policy makers, either directly or indirectly, and media events such as the high profile and TV coverage in Ecuador, in conjunction with the plenary ChagasEpiNet meeting in Quito. *Exploitation* is essentially as described above; some technological developments might eventually yield products for (non-profit) wide exploitation.