COMPOSITIONAL VARIATIONS ON 5'UTR OF TRANS-SIALIDASES AND POTENTIAL USE FOR TRYPANOSOMA CRUZI GENOTYPING

Myllena Fatima Alheiros Dias Melo[1], Franklin Souza da Silva[1], Otacilio Moreira[1], Christian Macagnan Probst[3]; Adeilton Brandão[2] and Constança Britto[1]

1 - Laboratório de Biologia Molecular e Doenças Endêmicas / Oswaldo Cruz Institute; 2 - Laboratório Interdisciplinar de Pesquisas Médicas / Oswaldo Cruz Institute; 3- Bioinformatics and Computational Biology Laboratory / Carlos Chagas Institute

FIOCRUZ

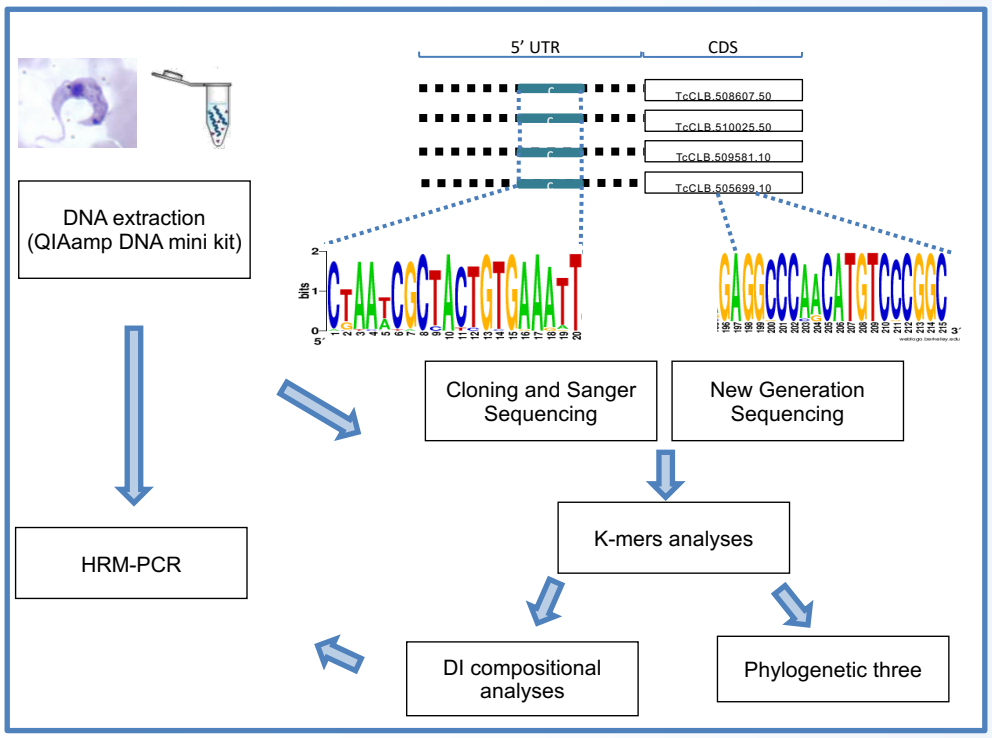
The work hypotheses: Is possible to characterize *T. cruzi* lineages through 5'UTR of trans-silaidase genes?

INTRODUCTION

Through Trypanosoma cruzi genome study, a number of genomic targets has been identified for genotyping purpose. Genotyping algorithms have been used to define the lineages, but interpreting results is somewhat subjective and time consuming. Besides the wide range of molecular targets already identified, the molecular diagnosis per se is neither capable of directly genotyping clinical samples, nor making a correlation between parasite and disease evolution. Therefore, in this work, we evaluated the nucleotide composition of trans-sialidase genes to characterize strains. Apart from more than 1,400 copies in the genome and parasite virulence relation, trans-sialidase may be a potential molecular target to be used for Chagas's disease molecular diagnoses. Different strains of *T. cruzi* appear to have unique sets of these genes, suggesting that TS gene families had evolved in lineage under certain immune pressure. Variations in this multigenic family, in special on 5'UTR (untranslated region) could determine biological strain profile.

METHODS

To assess the potential of a segment of trans-sialidase (TS), we designed oligonucleotides for a segment containing part of 5'UTR and a portion of TS coding region. Amplification of this segment followed by sequencing allowed us to carry out the nucleotide compositional analysis, which revealed important motif on 5'UTR of TS. New oligonucleotides were designed for these regions and submitted to HRM-COLD-PCR.



RESULTS AND DISCUSIONS

1) PCR amplification of a 200 bp fragment from *T. cruzi* strains

Oligonucleotides were designed to hybridize a region containing part of 5' UTR and an initial segment of an ORF of trans-sialidases, which is localized after the first ATG defined by genomic annotation. PCR amplification resulted in a 200 bp fragment detected in samples which carried the *T. cruzi* DNA, as well as *T. cruzi* bat and *T. cruzi* marinkellei. Furthermore, no cross-reactions with *Leishmania* sp, *T. rangeli* and *Toxoplasma gondii* DNA were observed. A fragment from all *T. cruzi* strains was cloned and sequenced by Sanger. This process resulted in a total of 487 sequences, which showed more than 95% of cover and identity with 32 CL Brener reference genes (NCBI). Also, they were distributed in twelve different chromosomes (Table 1).

Table 1. Sequencing analysis of PCR products of *T. cruzi* strains

Strains	DTU	Number of Sequence (sanger)	Median of sequence lengths	Number of exclusive fragments with Kmer 20	Number of Exclusive DTU sequences
Colombian	Tc I	77	198 bp	26	5
DM28c	Tc I	35			
Tulahuen cl0	Tc I	16			
D8	Tc I	25			
Y	Tc II	63	198 bp	14	3
3663	Tc III	49	188 pb	12	5
4167	Tc IV	50	200 pb	225	13
LL014	Tc V	129	200 pb	14	4
CL	Tc VI	43	198 bp	24	1
Total		487		-	31

2) Identification of dinucleotide composition on a conserved regions of 5' UTR

Through k-mers analysis, we identified specific polymorphisms, which allowed us to select sequences which are DTU exclusive. As a result of alignment, we observed two conserved blocks in these sequences (first block: ATGCTCTCACGTGTT; second block: CCGGATCCAGCG) (Fig 1). Thus, we analyzed the dinucleotides (DI) composition per section (separated by conserved blocks) in order to identify potential compositional variation between *T. cruzi* lineages. There was a higher frequency of CA and AC in section 1 and AG, GA and GG in section 3 ($p<0.05$, using Tukey's multiple comparisons test). Moreover, there was an absence of TT, TA and CT in section 3. Section 2 showed a more homogeneous DI frequency compared with sections 1 and 3.

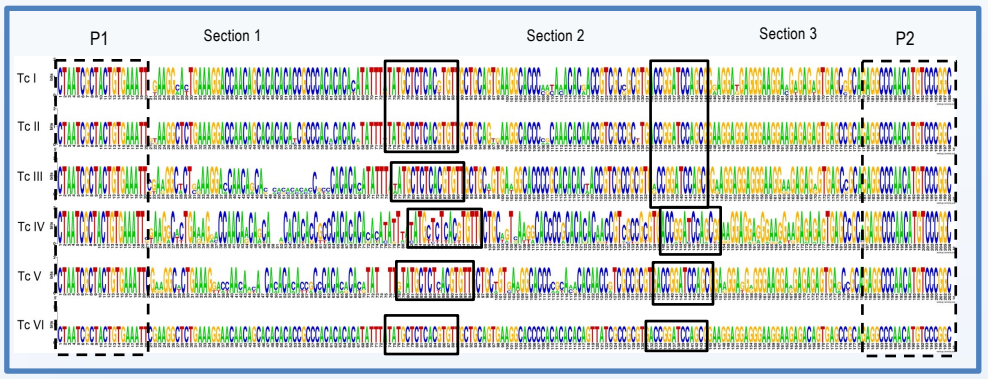
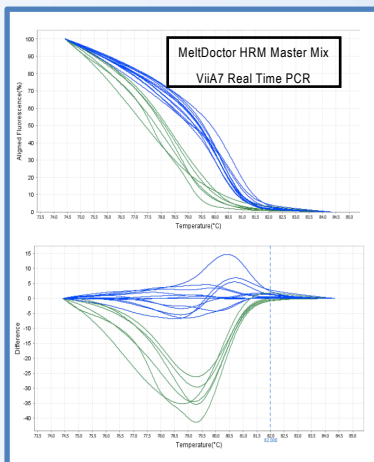


Figure 1. Distribution of the DI frequency values per section and between *T. cruzi* lineages.

3) High Resolution Melting: molecular signatures for Trypanosoma cruzi identification

CO amplification at Lower Denaturation temperature (COLD) has been used to detect low frequency mutant sequences for cancer diagnosis, as well as to minimize heteroduplex formation during pre-dissociation of fragments in heterogeneous samples. Amplification of section 2, which enhanced exclusive segments of TS from different DTUs, showed more potential to group strains / clones that are genetically close: DM28c, D8, Colombiana and 3663 (TcI and TcIII) are together in the same variant (green), while Y, PR1281, LL014, SO3, CL and 4167, (TcII, TcV, TcVI and TcIV) are in another variant (blue) (Fig 2). In addition, the amount of conserved nucleotide was 53% and 86% for variant 1 (blue) and 2 (green), respectively, on sequences from section 2.



Variant	Complete fragment (200bp)	Section 1	Section 2	Sections 2 and 3
1	Y, LL014, CL, 4167, D8	CL, Y, 3663, 4167, D8, LL014	CL, Y, 4167, LL014, PR1281, SO3	CL, Y, 3663, LL014, D8, SO3, Colombiana, PR1281, DM28c
2	SO3, Colombiana, PR1281, LL014, DM28c, 3663	DM28c, Colombiana, PR1281	DM28c, 3663, D8, Colombiana	CL, Y, 3663, LL014, D8, PR1281
3	DM28c, 3663, CL, 4167, D8	SO3	-	4167, DM28c

Human infection | High virulence

Wild animal infection | Low virulence

CONCLUSION

Moreover, genotyping algorithms have been used to define *T. cruzi* lineages is well defined, the 5' UTR of TS showed potential to group genetically closed strains. Amplification of a segment on 5'UTR between 2 conserved blocks showed an efficiency to group strains in two variants, suggesting an association with transmission cycles, as well as strains virulence. The COLD-PCR used on cancer diagnoses should be an approach to amplify low frequency TS copies. In addition, his segment is a potential target to be used as a complementary test for the molecular diagnoses of Chagas disease. Finally, 5' UTR of TS could be apply for large-scale sequencing in order to characterize *T. cruzi* strains in six DTUs using only one gene.

Financial support: Capes, CNPq, FAPERJ, Fiocruz