African Journal of Biotechnology

Volume 16 Number 25, 21 June, 2017

ISSN 1684-5315

ABOUT AJB

The African Journal of Biotechnology (AJB) (ISSN 1684-5315) is published weekly (one volume per year) by Academic Journals.

African Journal of Biotechnology (AJB), a new broad-based journal, is an open access journal that was founded on two key tenets: To publish the most exciting research in all areas of applied biochemistry, industrial microbiology, molecular biology, genomics and proteomics, food and agricultural technologies, and metabolic engineering. Secondly, to provide the most rapid turn-around time possible for reviewing and publishing, and to disseminate the articles freely for teaching and reference purposes. All articles published in AJB are peer-reviewed.

Contact Us

Editorial Office: ajb@academicjournals.org

Help Desk: helpdesk@academicjournals.org

Website: http://www.academicjournals.org/journal/AJB

Submit manuscript online http://ms.academicjournals.me/

Editorial Board

African Journal of Biotechnology

Table of Content: Volume 16 Number 25 21 June, 2017

ARTICLES

academicJournals

Vol. 16(25), pp. 1366-1378, 21 June, 2017
DOI: 10.5897/AJB2017.16021
Article Number: 6383CB064862
ISSN 1684-5315

http://www.academicjournals.org/AJB

African Journal of Biotechnology

Full Length Research Paper

Kinetic characterization and molecular modeling of trehalose-6-phosphate phosphatase from *Anopheles gambiae* and expressed in *Pichia pastoris*

Marcos Cézar Fernandes Pessoa[1], Érika Izumi[2], Sandra Patrícia Zanotto[1] and Spartaco Astolfi-Filho[1]*

[1]Laboratory of DNA Technology, Biotechnology Department, Multidisciplinary Support Center, Federal University of Amazonas, Manaus, Amazonas, Brazil.
[2]Federal University of Technology, Curitiba, Paraná, Brazil.

Received 4 April, 2017; Accepted 14 June, 2017

Trehalose-6-phosphate phosphatase (TPP) is one of the primary enzymes involved in the synthesis of trehalose, the main sugar found in insect hemolymph. In the present study, we report for the first time heterologous expression in *Pichia pastoris*, characterization and homology modeling of TPP from *Anopheles gambiae* mosquito. Purified TPP recombinant exhibited a molecular weight of approximately 36 kDa with optimum pH of 8.0, optimum temperature of 38°C, and the K_M was 3.19 ± 0.10 mM. Inhibition tests revealed that $CaCl_2$ and $Ca(NO_3)_2$ at 5 and 25 mM, respectively, were effective inhibitors of TPP activity. Homology studies and molecular modeling indicated high similarity of the TPP tridimensional structure with TPP enzymes deposited at a data bank and these homology studies confirmed that it is a trehalose phosphatase with conserved motifs of the superfamily haloacid dehydrogenase. These results may provide support for the discovering of TPP inhibitors to be used for development of insecticides to control mosquito vectors.

Key words: Trehalose, trehalose-6-phosphate phosphatase, expression, enzyme characterization, homology modeling.

INTRODUCTION

Malaria is a potentially serious infectious disease caused by unicellular protozoans of the genus *Plasmodium* species and is transmitted to vertebrate hosts through bites from females of the mosquito *Anopheles* species (Forattini, 2002; WHO, 2016). Human malaria is a parasitic disease with medical, social and economic importance, and it constitutes one of the most important public health problems worldwide. Malaria is endemic in 91 countries, with 212 million estimated cases and 429 000 estimated deaths from malaria globally in 2015 (WHO, 2016). The vast majority of malaria deaths occur in Africa, especially among children (Garcia, 2010), and the mosquito *Anopheles gambiae* Giles, belonging to the subgenus *Cellia*, is the primary human malaria vector.

*Corresponding author. E-mail: spartaco.biotec@gmail.com. Tel: + 55 92 3305 3240.

Resistance in the mosquito vector to at least one of the insecticides used for malaria control has been detected in 60 countries (WHO, 2016). Thus, studies on the metabolism of the insect vector and main enzymes involved in different metabolic pathways are therefore important for collecting data on the insect's physiology, which can be used in the research and design of insect control strategies.

Trehalose is a non-reducing glucose disaccharide and the primary carbohydrate found in insect hemolymph, thus, its metabolism is a potential target for the rational development of insecticides. This disaccharide serves as a circulating energy source that is similar to glucose in vertebrate blood (Klowden, 2013), and it is synthesized in the insect fat body through the combined action of two enzymes, trehalose-6-phosphate synthase (TPS) and trehalose-6-phosphate phosphatase (TPP) (Karthik et al., 2011). These enzymes have been detected in insects and other invertebrate animal classes but not in humans. The regulation of the trehalose metabolic pathway in insects is well known (Thompson, 2003).

The trehalose biosynthetic pathway was first described in *Sacharomyces cerevisiae* (Cabib and Leloir, 1958). At the insect fat body, trehalose-6-phosphate synthase catalyzes the formation of a trehalose α-1, α-1 glycosidic bond by the condensation of glucose-6-phosphate and glucose from uridine diphosphate glucose, forming trehalose-6-phosphate (T6P). The second reaction is catalyzed by TPP and involves the cleavage of phosphate ester bonds and the release of inorganic phosphate and free trehalose to the hemolymph (Eastmond et al., 2002; Matula et al., 1971). TPP (EC 3.1.3.12) is a member of the superfamily haloacid dehalogenase (HAD) (Collet et al., 1998) and exhibits significant sequence homology with several phosphatases, P-type ATPases (Koonin and Tatusov, 1994) and other members of the family (Rao et al., 2006). TPP displays specific activity for substrate T6P, resulting in trehalose formation in the presence of water, which is characteristic of enzymes of the hydrolase family with phosphatase activity (Elbein, 2009; Klutts et al., 2003; Kormish and McGhee, 2005; Rao et al., 2006).

Gene silencing techniques applied do TPS gene from *Aedes aegypti* (Vaidyanathan et al., 2015) and suppression of activity of trehalase from *Nilaparvata lugens* by inhibitors (Tang et al., 2017) have shown that this pathway is fundamental in the maintenance and survival of insects. In the present study, we report on the expression, characterization and homology modeling of TPP from *A. gambiae.* These results can support the development of new approaches that using the TPP as a molecular target to discovery novel insecticides for use in mosquito vectors control.

MATERIALS AND METHODS

TPP synthetic gene design

The nucleotide sequence of the TPP gene structural region was obtained from the *A. gambiae* genome data bank, which is available at the Kyoto Encyclopedia of Genes and Genomes (KEGG) website (www.genome.jp/kegg/). The DNA sequence was flanked with restriction sites for specific recognition by the endonucleases *Eco*RI and *Not*I, which enabled subcloning of the sequence into the expression and secretion vector for *Pichia pastoris*, pPIC9. A sequence coding a six-histidine tail (*his*^{6}*-tag*) was inserted at the 3' end for purification and identify the confirmation of the expressed protein by immunological assays. The nucleotide sequence was chemically synthesized by GenOne (Rio de Janeiro, RJ, Brazil) and cloned into the vector pBSK. The vector containing the synthetic *TPP* gene was named pBSK-T.

The structural sequence of the *TPP* gene optimized for expression in *P. pastoris* was constructed using the bioinformatics tools KEGG, BLAST, ExPasy, UniProt, CBS Prediction Servers, NEBcutter V 2.0 and Clustal; these tools are available at specialized websites. The theoretical properties of the translated recombinant protein of the *TPP* gene were determined using the ExPASy ProtParam Tool (http://web.expasy.org/protparam/) and the NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/) and NetOGlyc 3.1 Server (http://www.cbs.dtu.dk/services/NetOGlyc-3.1/) bioinformatic tools.

Cloning and expression

To construct the TPP expression and secretion vector for *P. pastoris*, the synthetic *TPP* coding sequence was freed from the pBSK-T cloning vector by double digestion with *Eco*RI and *Not*I according to the enzyme manufacturer's recommendations (New England Biolabs, Ipswich, MA, USA). The fragment was then purified from agarose gel using an Illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare, Little Chalfont, UK) and ligated into the *Eco*RI and *Not*I sites of the multiple cloning site of the pPIC9 vector. The recombinant plasmid was named pPIC-TPP and was used in the transformation of *Escherichia coli* DH10B electrocompetent cells. Plasmid DNA extracted from the *E. coli* DH10B host was subjected to restriction analysis to confirm the correct construction of the recombinant plasmid pPIC-TPP.

The plasmid pPIC-TPP was digested with *Bgl*II for 2 h at 37°C to free the expression cassette containing the TPP coding sequence. The digestion product was used to transform the *P. pastoris* GS115 strain via electroporation, and recombinant clones were selected via auxotroph complementation in minimal dextrose (MD) medium without the addition of histidine. Plates were incubated for 72 h at 30°C.

Heterologous protein expression was induced with 25 ml of buffered glycerol-complex medium (BMGY) and incubated at 30°C under constant stirring at 250 rpm for approximately 24 h. The cells were collected following centrifugation at 4000 rpm for 10 min at 4°C and then resuspended in 50 ml of buffered methanol complex medium (BMMY). The promoter AOX1, which modulates the transcription of the heterologous gene inserted into the host genome, was induced by adding 250 µl of methanol (Pro Analyse) to the cell culture every 24 h to a 0.5% final concentration (v/v). *TPP* gene expression was induced for 72 h at 30°C under constant stirring at 180 rpm.

Western blot analysis

The recombinant clone expressing TPP was selected via colony blot immunoassay, which was performed in plates according to the QIAexpressionist manual (QIAGEN, Hilden, DE) for transformant analysis. The expression of recombinant TPP was confirmed through western blot analysis. Following sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), the protein was transferred to a nitrocellulose membrane through a semi-dry

transfer system (Gravel, 2002) performed with a Trans-Blot® SD Semi-Dry Transf Cell (Bio-Rad, Hercules, CA, USA) for 1 h under the following conditions: 10 volts, 0.13 A, 1 W. Following the transfer, the nitrocellulose membrane containing the immobilized protein was treated with the antibodies and reagents contained in the WesternBreeze® Chromogenic Western Blot Immunodetection Kit (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions.

Enzyme purification

The supernatant containing the recombinant protein was filtered using a retention membrane with 0.45 µm diameter pores. TPP purification was performed using an ÄKTA purifier (GE Healthcare) coupled to a HisTrap™ 1-ml column (GE Healthcare). The column was pre-equilibrated with 2 ml of binding buffer (20 mM NaH_2PO_4•H_2O; 0.5 M NaCl; pH 7.4). The sample was eluted using an elution buffer (20 mM NaH_2PO_4•H_2O; 0.5 M NaCl; pH 7.4) with an imidazole gradient up to 0.5 M. The fractions were collected at room temperature at a flow rate of 1 ml/min, and absorbance was read at 280 nm. The fractions containing the recombinant protein were analyzed via an SDS-PAGE gel (Laemmli, 1970) stained with Coomassie Brilliant Blue according to the methods of Oakley et al. (1980).

Purified recombinant protein was quantified using a Pierce™ BCA Protein Assay kit (Thermo Scientific, Waltham, MA, USA) according to the manufacturer's recommendations. The assay was performed in triplicate using bovine serum albumin (BSA) as the standard.

Enzyme characterization

Enzyme characterization was performed using 200 ng/µl recombinant enzyme and 1 mM T6P. The assay for TPP activity was measured by determining the release of inorganic phosphate from T6P according methodology descripted by Edavana et al. (2004) and Klutts et al. (2003).

pH effect

The pH effect on the enzyme activity was determined at 37°C and 10 min incubation using the following buffers: 2-ethanesulfonic acid (MES) (pH 5.0 to 7.0), Tris-HCl (pH 7.0 to 9.0) and glycine (pH 8.0 to 11.0).

Temperature effect

The optimum temperature was determined by initially incubating the reaction system over a temperature range between 30 and 80°C and then over a temperature range between 30 and 40°C.

Determination of kinetic constants

Enzyme activity was determined according to a modification of the method by Klutts et al. (2003), which used 1 to 8 mM T6P (Sigma-Aldrich, St. Louis, MO, USA), 2 mM $MgCl_2$, 50 mM Tris-HCl buffer and 1 µl of recombinant enzyme. The reaction was performed for 10 min, and the Michaelis-Menten curve and Lineweaver-Burk plot were obtained.

Effect of chemical inhibitors on TPP activity

TPP inhibition was tested using ethylenediaminetetraacetic acid (EDTA), NaF, $CaCl_2$ and $Ca(NO_3)_2$ at concentrations of 5 and 25 mM and final volume of 50 µl. The reaction was performed for 10 min.

Homology modeling of related TPP enzymes

The identity of the TPP sequence was confirmed through an alignment with TPP from other mosquito species using a BLAST search on the UniProtKB website (http://www.uniprot.org/). Sequences for each vector mosquito were saved in FASTA format. A global sequence alignment of TPP from *A. gambiae* with TPP sequences from other species of vector mosquitoes was then performed, and these sequences were compared with the template sequence from *Thermoplasma acidophilum* (PDB ID: 1U02).

The TPP sequences were submitted to the EMBL-EBI server (http://www.ebi.ac.uk/Tools/msa/clustalw2/) for multiple alignment using the software Clustal Omega 1.2.1 (Sievers et al., 2011). The TPP amino acid sequence was then submitted to the Phyre^2 server (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) (Kelley and Sternberg, 2009), and its secondary and tertiary structure was predicted and then compared with tridimensional protein structures deposited at the Protein Data Bank. The theoretical tridimensional structure of TPP was based on the template structure from *T. acidophilum* (1U02). The stereochemical quality of the structure was examined using PROCHECK (Laskowski et al., 1993) and subsequently refined using the software ModRefiner (Xu and Zhang, 2011). The protein was visualized using the software PyMOL (DeLano, 2006).

Modeling the TPP-T6P complex

The molecular docking of the substrate to the TPP active site was modeled using AutoDock 4.2.6 (Morris et al., 1998), a Lamarckian genetic algorithm (Solis and Wets, 1981) and the software PyRx 0.8 (Wolf, 2009). The substrate molecule was obtained from the PubChem Compound Database (Bolton et al., 2008) minimized and converted to PDBQT format using PyRx. The grid box that defines the size of the binding pocket was created using the software AutoGrid (Morris et al., 1998). The grid box was centered at coordinates x=20.8224, y=22.0076 and z=16.1604 Å and the spacing between grid points was 0.375 Å. The AutoDock results were visualized using PyRx, and the most favorable ligand-binding pose was selected based on its conformation and binding energy. The selected docking was analyzed using the Discovery Studio 4.0 Visualizer software (Accelrys, 2013), which was also used to identify the number of hydrogen bonds and other enzyme-substrate intermolecular interactions.

Statistical analysis

The enzyme characterization data (mean + standard error [SE]) were analyzed by performing analysis of variance (ANOVA) using the software GraphPad Prism 5.04 for Windows. The inhibition test data were subjected to Bonferroni correction at $p<0.05$. All of the assays were performed in triplicate. The Bonferroni correction is a standard test of the program used for this analysis type.

RESULTS AND DISCUSSION

The enzyme TPP is present in bacteria, fungi, plants and invertebrates but not in mammals. Its catalytic product (trehalose) is used by these organisms as an energy reserve and for protection against environmental

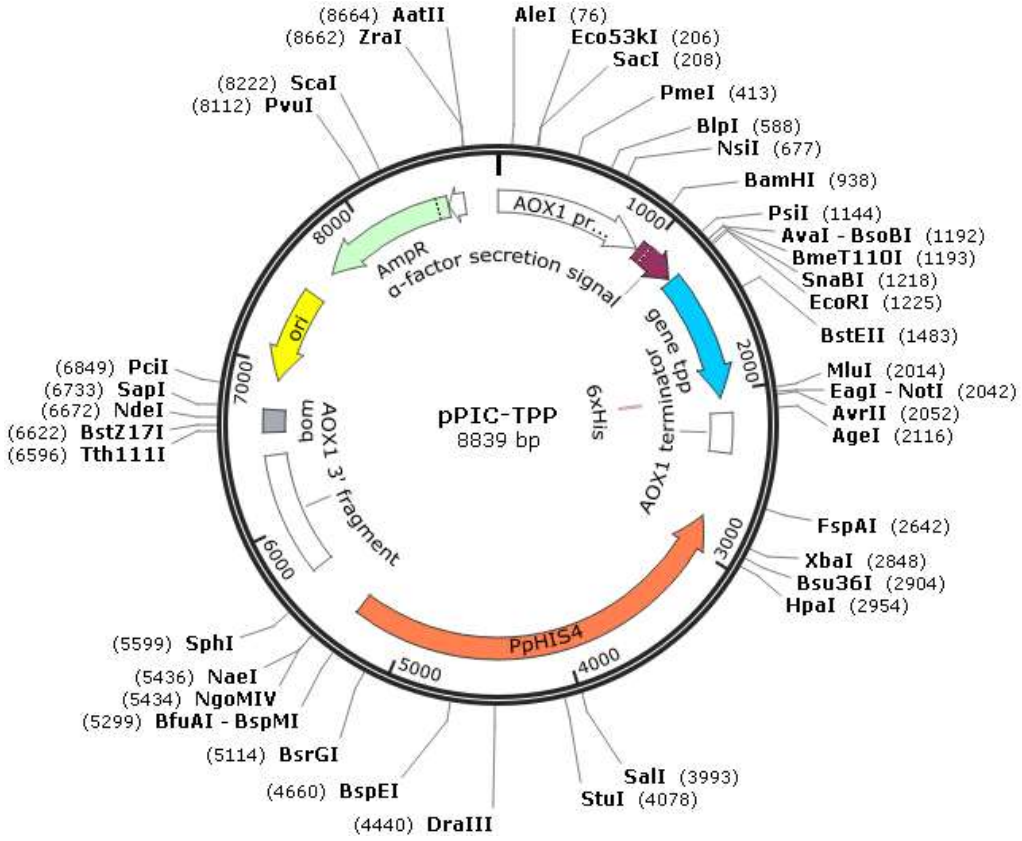

Figure 1. Physical map of the pPIC-TPP recombinant plasmid showing the insertion of the *TPP* gene (in blue) into the vector pPIC9, which is flanked by restriction enzymes *Eco*RI and *Not*I and located after the coding region for the α-factor signal peptide for secretion. The physical map was constructed using the software SnapGene® Viewer 2.4.2.

stresses, such as oxidative and osmotic stress, freezing, desiccation and anhydrobiosis (Behm, 1997; Elbein et al., 2003).

TPP is important for the metabolism and physiology of the mosquito *A. gambiae*, which is an important vector of *Plasmodium falciparum*, the parasite responsible for the most serious form of human malaria in Africa. The nucleotide and amino acid sequences of the *TPP* gene were obtained from the KEGG website and can be found at the following accession: *A. gambiae* (mosquito): AgaP_AGAP008225-PA. The information on this gene available at KEGG are in accordance with the National Center for Biotechnology Information (NCBI) and UniProtKB (Protein Knowledgebase - AgTpp AGAP008225/AgaP_AGAP008225) databases. The *TPP* gene was selected because it codes for a protein that is involved in the formation of disaccharide trehalose in the insect's carbohydrate metabolic pathway. TPP was chosen for cloning and expression because its catalytic product is trehalose, which is not essential for mammal cells, and not glucose, which is the major carbohydrate required for numerous mammalian metabolic pathways (Karthik et al., 2011; Klutts et al., 2003; Kormish and McGhee, 2005; Kushwaha et al., 2011).

After codon optimization and chemical synthesis, the *TPP* gene structural region was cloned into the plasmid vector pPIC9 under control of the AOX *P. pastoris* strong promoter. The physical map of the pPIC-TPP recombinant plasmid resulting from insertion of the *TPP* gene into the expression/secretion vector is presented in Figure 1. The TPP coding sequence is located downstream from the coding sequence for the *S. cerevisiae* α-factor signal peptide, which ensures TPP secretion.

Recombinant clones were obtained through the transformation of *E. coli* DH10B with the ligation product of the *TPP* gene structural region to the vector pPIC9. A transformant clone was selected for subsequent plasmid DNA extraction, and confirmation of the correct construction of the pPIC-TPP recombinant plasmid was performed through restriction analysis. The resulting electrophoresis profile revealed that the freed fragments correspond to the size of the vector pPIC9 (8023 bp) and the *TPP* gene structural region (830 bp), confirming that the structure of the pPIC-TPP recombinant vector was correct (Figure 2A).

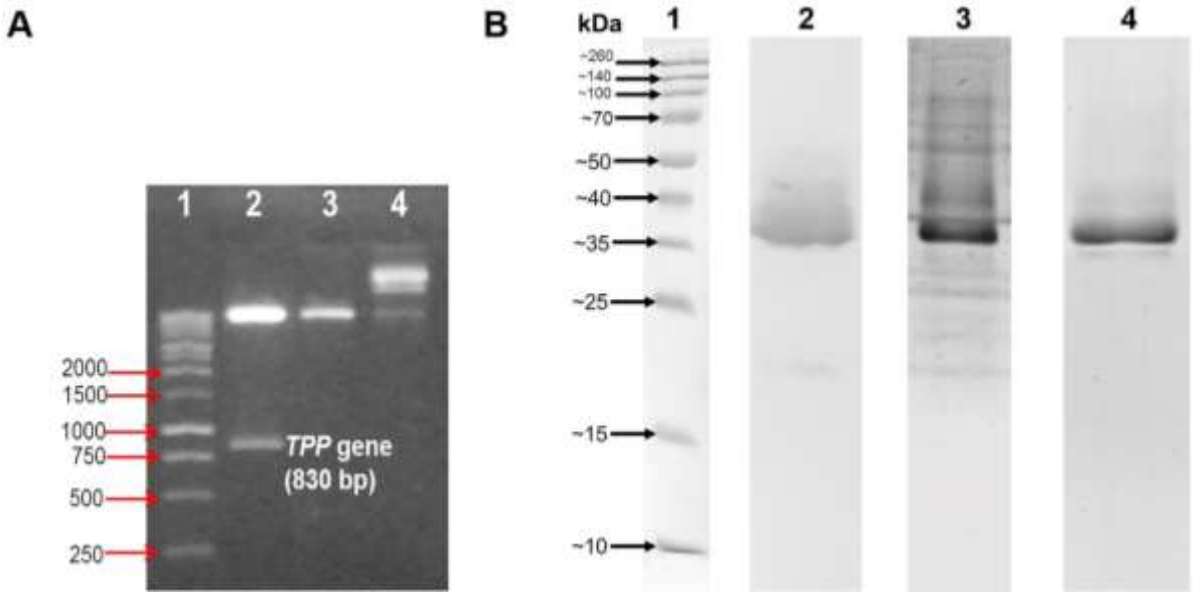

Figure 2. Cloning, expression and purification of TPP from *A. gambiae.* **(A)** Restriction analysis of the pPIC-TPP recombinant plasmid. Lane 1, Marker GeneRuler™ 1 Kb DNA Ladder (Fermentas); lane 2, pPIC-TPP digested with *Eco*RI and *Not*I; lane 3, pPIC9 without insert digested with *Eco*RI and *Not*I; lane 4, pPIC9 not digested. **(B)** Electrophoretic profiles of recombinant TPP. Lane 1, Marker Spectra Multicolor Broad Range Protein Ladder (Thermo Scientific); kDa molecular weights marker varying between 10 and 260 kDa; lane 2, western blotting of recombinant TPP induced for 72 h in BMMY medium; lane 3, electrophoretic profile (15% SDS-PAGE gel) of recombinant TPP; lane 4, electrophoretic profile (15% SDS-PAGE gel) of TPP purified by affinity chromatography.

The expression/secretion of the heterologous protein by a selected clone of *P. pastoris* after incubation for 72 h was analyzed by western blotting of the supernatant using the anti-His^{6}-tag antibody. The results indicated that TPP was efficiently expressed by the *P. pastoris* expression system (Figure 2B, lane 2).

Following confirmation of TPP expression and secretion in the supernatant, recombinant TPP was newly induced. The culture supernatant was analyzed by a 15% SDS-PAGE gel. The TPP recombinant protein was larger than expected for the native protein (Figure 2B, lane 3). In addition, TPP was secreted in large amounts in the supernatant and constituted the largest portion of proteins produced by the yeast, which was observed in the supernatant. A bioinformatics analysis was performed using *ExPASy* and *ProtParam* and estimated a 29.3 kDa molecular weight for non-glycosylated TPP. However, the SDS-PAGE band corresponding to recombinant TPP exhibited electrophoretic migration equivalent to 36 kDa. Because the TPP polypeptide sequence possesses two potential N-glycosylation sites and one O-glycosylation site and because glycosylation affects protein electrophoretic migration in an SDS-PAGE gel, it was inferred that the estimated weight of the recombinant TPP increased from 29.3 kDa to approximately 36 kDa as a result of glycosylation during post-translational processing in the secretory pathway. Recombinant proteins expressed in *P. pastoris* may be larger than expected and exhibit different sizes when analyzed by SDS-PAGE if they are glycosylated (Cregg et al., 2009) or hyperglycosylated (Cereghino and Cregg, 2000).

TPP was purified in one stage using Ni^{2+} resin affinity chromatography and was eluted with 300 mM imidazole upon application of an elution gradient. The purified recombinant protein was analyzed to SDS-PAGE (Figure 2B, lane 4). According to the electrophoretic profile, the recombinant TPP was eluted as a single band that migrated with an apparent 36 kDa molecular weight, which is in accordance with the results presented in Figure 2B.

Following TPP purification, its physicochemical properties were determined. The effects of pH, temperature and inhibitors on TPP activity were determined as well as kinetic constants. To our knowledge, there are no previous characterization studies of TPP from *A. gambiae*.

Recombinant TPP activity was tested in MES-Tris-Glycine buffer in a pH range from 5.0 to 11.0 in 1.0 unit intervals. The recombinant TPP activities at the different pH levels are presented in Figure 3A.

P. pastoris secreted TPP exhibited higher activity between pH 7.0 and 8.0, with a peak at pH 8.0, and it maintained approximately 25% activity at pH 5.0 and 9.0. The optimum TPP pH varies widely among organisms. Among insects, higher activity at pH 7.0 was observed for TPP from the cockroach *Periplaneta Americana* (Friedman

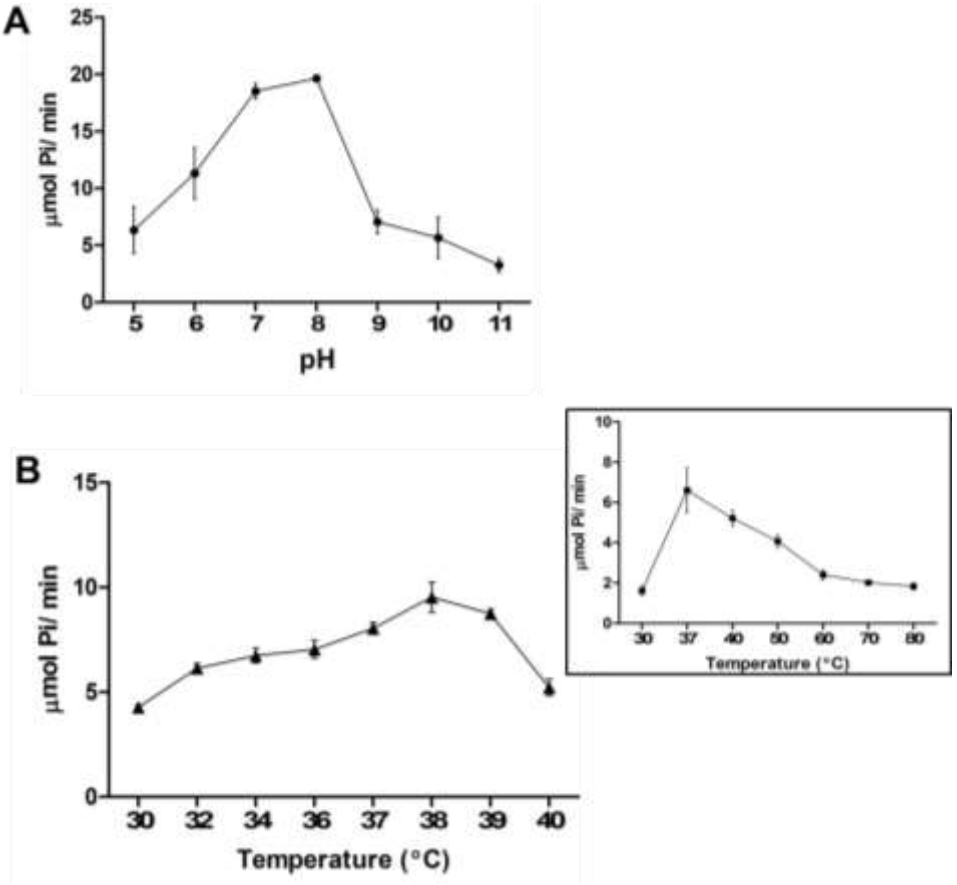

Figure 3. Characterization of recombinant TPP enzymatic parameters. (A) Determination of optimum pH: enzyme activity was measured in MES-Tris-Glycine buffer (50 mM) at different pH (5-11). Optimum pH for *A. gambiae* TPP activity was 8.0. (B) Determination of optimum temperature: enzyme activity was measured at different temperatures, and the enzyme activity over a wide temperature range (30 - 80°C) is presented in the right upper corner, with activity peak at 37°C. The optimum temperature over a narrow temperature range (30 - 40°C) is highlighted. Based on this range, the TPP optimum temperature, at which it exhibits maximum activity, was 38°C.

and Hsueh, 1979). TPP from the black blow fly (*Phormia regina*) also exhibited the highest activity at pH 7 and maintained activity between pH 6.0 and 8.5 (Friedman, 1966). Therefore, the optimum TPP pH does not present large variations among different classes of insects.

Regarding the temperature effect, purified TPP was first evaluated at a temperature range between 30 and 80°C in Tris-HCl buffer at pH 8.0 (Figure 3B, right upper corner), and its catalytic action was then evaluated at a narrower temperature range, between 30 and 40°C (Figure 3B, highlighted). Recombinant TPP exhibited a maximum activity of 9.53 µmol Pi/min at 38°C (Figure 3B). This temperature was therefore considered the optimum temperature for the activity of TPP from *A. gambiae*.

Friedman (1966) studied *P. regina* and reported that the highest TPP activity was at 46°C and that 85% activity was maintained at 57°C for 15 min. Friedman and Hsueh (1979) determined TPP activity from *P. americana* at 32°C. These results indicate that TPP exhibits a wide variation in optimum temperature profiles.

Enzyme activity was tested at different T6P substrate concentrations (Figure 4). Increasing the substrate concentration only increases the reaction velocity up to the T6P concentration that produces the maximum TPP catalytic capacity (maximal velocity).

The substrate concentration curve (T6P [mM]) versus the initial velocity (V_0) is shaped as the perfect rectangular hyperbola (Figure 4) described by the Henri-Michaelis-Menten equation (Segel, 1979). At low substrate concentrations, the initial reaction velocity was approximately proportional to the T6P concentration, indicating the first order kinetic region. As the concentration of the substrate increases, the initial velocity of the reaction is no longer proportional to the concentration of the substrate.

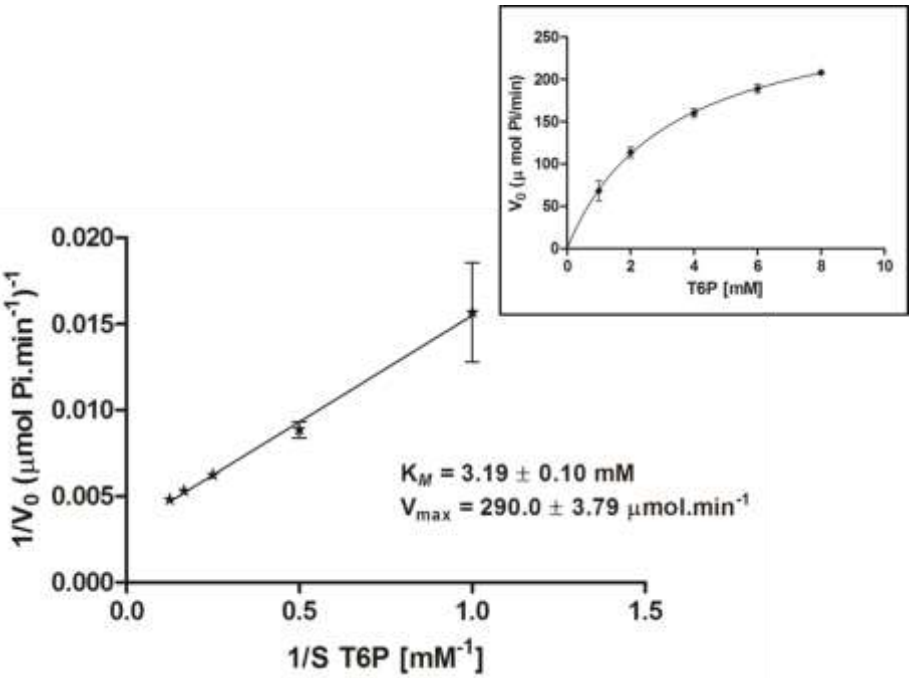

Figure 4. Kinetic constants of recombinant TPP from *A. gambiae*. K_M and V*max* were determined in triplicate using non-linear regression. The Michaelis-Menten curve was obtained with substrate concentrations between 1 - 8 mM and is presented at the right upper corner. The Lineweaver-Burk plot is highlighted.

The substrate concentration necessary to reach half-maximal velocity (½ *Vmax*) corresponds to K_M, the Michaelis-Menten constant (Figure 4) (Campbell and Farrell, 2007). The K_M value is important for the characterization of enzyme kinetics because different enzymes may have the same maximal initial velocity but generally will not have the same K_M, which indicates the difference in the amount of substrate required to saturate the enzyme (Vieira, 2003).

The recombinant TPP *Vmax* was 290.0 ± 3.79 µmol.min^{-1}, and K_M was 3.19 ± 0.10 mM, which is on the same order of magnitude of most of the reported TPPs: 2.5 ± 0.1 mM (Seo et al., 2000), 1.5 mM (Silva et al., 2005), 6.1 ± 1.5 mM (Nobre et al., 2008), 0.42 mM (Kushwaha et al., 2011), 0.39 ± 0.16 mM (Li et al., 2012), 0.36 ± 0.06 mM (Farelli et al., 2014), 0.87± 0.06 mM (Miao et al., 2016), 0.48 mM (Shan et al. 2016), 0.23 ± 0.07, 0.50 ± 0.10, 0.69 ± 0.07 and 0.31 ± 0.04 mM (Liu et al., 2017a).

Inhibition tests were performed to search for organic or inorganic chemical inhibitors of the recombinant TPP. The results of chemical inhibition tests on TPP activity at inhibitor concentrations of 5 and 25 mM are presented in Figure 5.

All of the tested chemical compounds inhibited TPP activity at 25 mM. At 5 mM, only two of the tested compounds ($CaCl_2$ and $Ca(NO_3)_2$) significantly inhibited TPP activity (Figure 5).

An ANOVA with Bonferroni correction performed on the results presented in Figure 5 confirmed that all of the chemical inhibitors tested at 25 mM resulted in significant inhibition (p<0.001; <0.0001) compared with that of the control, indicating that these compounds effectively inhibited TPP activity at this concentration. However, only two of the tested compounds at 5 mM ($CaCl_2$ and $Ca(NO_3)_2$) effectively inhibited recombinant TPP (p<0.001 and p<0.05, respectively).

These results indicate that compounds containing the Ca^{2+} divalent cation more effectively inhibit TPP, which is a result of the ion competing with Mg^{2+} for the enzyme active site, particularly for the substrate phosphate group, which characterizes competitive inhibition. Compounds containing Ca^{2+} ions have an inhibiting effect on phosphatases (Mamedov et al., 2001; Peeraer et al., 2004).

EDTA and NaF are considered non-competitive inhibitors because their inhibition depends on the concentration of these compounds but no concentrations of substrate or cofactor (Matula et al., 1971).

EDTA was reported to inhibit TPP from the bacteria *Mycobacterium smegmatis* (Matula et al., 1971) and the nematode *Brugia malayi* (Kushwaha et al., 2011). EDTA inhibition is caused by its action as a chelating agent because it forms stable complexes with several metal ions, including Mg^{2+} (the TPP cofactor) at pH values higher than 7 (Holleman and Wiberg, 2001). EDTA

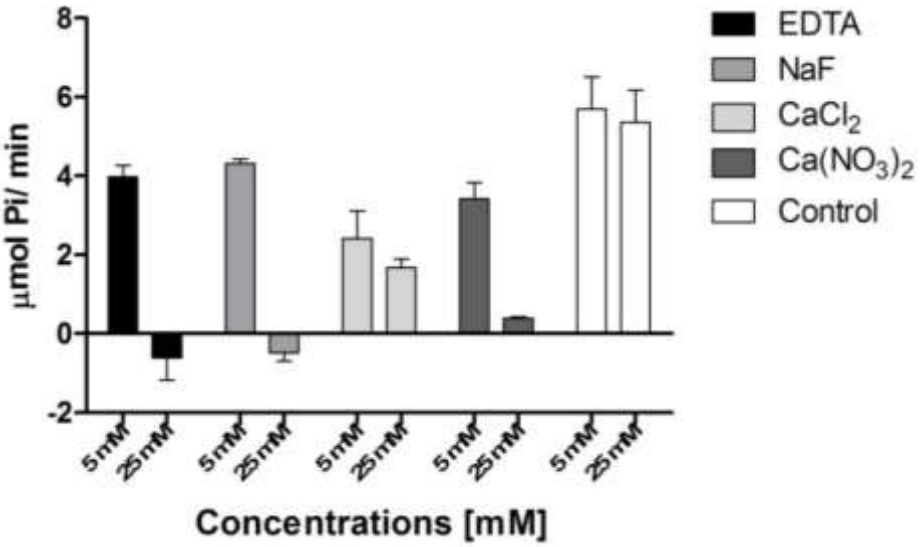

Figure 5. Effect of chemical inhibitors at 5 and 25 mM on recombinant TPP activity.

treatment of TPP from *T. acidophilum* resulted in the loss of its activity, suggesting that TPP is Mg^{2+} dependent (Rao et al., 2006). EDTA may therefore render the enzyme non-functional.

However, the inhibitory effect of NaF is caused by its action as a non-competitive phosphatase inhibitor (Matula et al., 1971). Friedman (1966) also observed that NaF inhibited TPP produced by *P. regina* by 70% when applied at 25 mM, which is the same concentration that inactivated recombinant TPP in the present study.

A standard BLASTP search at the UniProtKB web server identified several protein sequences with significant similarity to TPP from *A. gambiae*. The best hits were TPP from vector mosquitoes such as *Anopheles darlingi* (W5JSV8_ANODA), *Anopheles aquasalis* (T1E9M2_ANOAQ), *Culex quinquefasciatus* (B0WQL2_CULQU), *Aedes albopictus* (A0A023EMC7_AEDAL) and *Aedes aegypti* (Q16S69_AEDAE). The *A. gambiae* native enzyme (Q7PJ67_ANOGA) sequence was also considered in this search for comparison purposes. In the BLAST analysis, the trehalose phosphatase amino acid sequence showed 89% similarity to the TPP from *A. darlingi*, 88% similarity to *A. aquasalis*, 72% similarity to *C. quinquefasciatus*, 67% similarity to *A. albopictus* and 65% similarity to *A. aegypti*. These mosquitoes are important vectors of diseases transmitted by parasites and viruses, with *A. darlingi* the main vector of malaria in Neotropical regions, *A. aquasalis* a malaria vector in coastal areas of South and Central America, *C. quinquefasciatus* a vector for bancroftian filariasis, Oropouche virus and Nile fever, *A. albopictus* a vector for chikungunya virus and dengue and *A. aegypti* a vector for yellow fever, chikungunya, dengue and zyka. The sequences of TPP from *A. gambiae* and the remaining vector mosquitoes were subjected to multiple alignments using Clustal Omega from the European Molecular Biology Laboratory - European Bioinformatics Institute (EMBL-EBI) site (Figure 6).

TPP contains highly conserved motifs that are characteristic of the HAD superfamily (Figure 6). The motif I sequence has the DxD signature. The carboxylate group of the first aspartate (Asp) and carboxyl radical of the second Asp coordinate the cofactor Mg^{2+}. In addition, the first Asp of motif I acts as a nucleophile and forms an aspartyl intermediate during catalysis (Baker et al., 1998; Collet et al., 1997; Morais et al., 2000). Motif II is characterized by a highly conserved threonine or serine amino acid (Burroughs et al., 2006). Motif III is centered on a conserved lysine that occurs around the N-terminal portion of the α-helix region. Motifs II and III contribute to the stability of hydrolysis intermediaries. Motif IV is characterized by a conserved acidic residues. The acidic residues of motif IV typically exhibit one of three signatures: DD, GDxxxD or GDxxxxD (where x is any amino acid). For TPP from *A. gambiae*, the signature of motif IV is represented by GDxxxD acidic residues. Motifs I-IV are spatially arranged around the binding cleft at the end of the C-terminal region of the central sheet strands that form the active site of the HAD superfamily (Burroughs et al., 2006).

The multiple sequence alignment using Clustal Omega showed high similarity between TPP from different vector insects. However, the amino acids from vector insect proteins exhibited low similarity with the 1U02 template sequence; the amino acid composition of the motif regions from the template sequence and vector mosquitoes were similar. The catalytic site residues of TPP from *T. acidophilum* (1U02) have also been found in insect TPP. The only disparities were the residues of

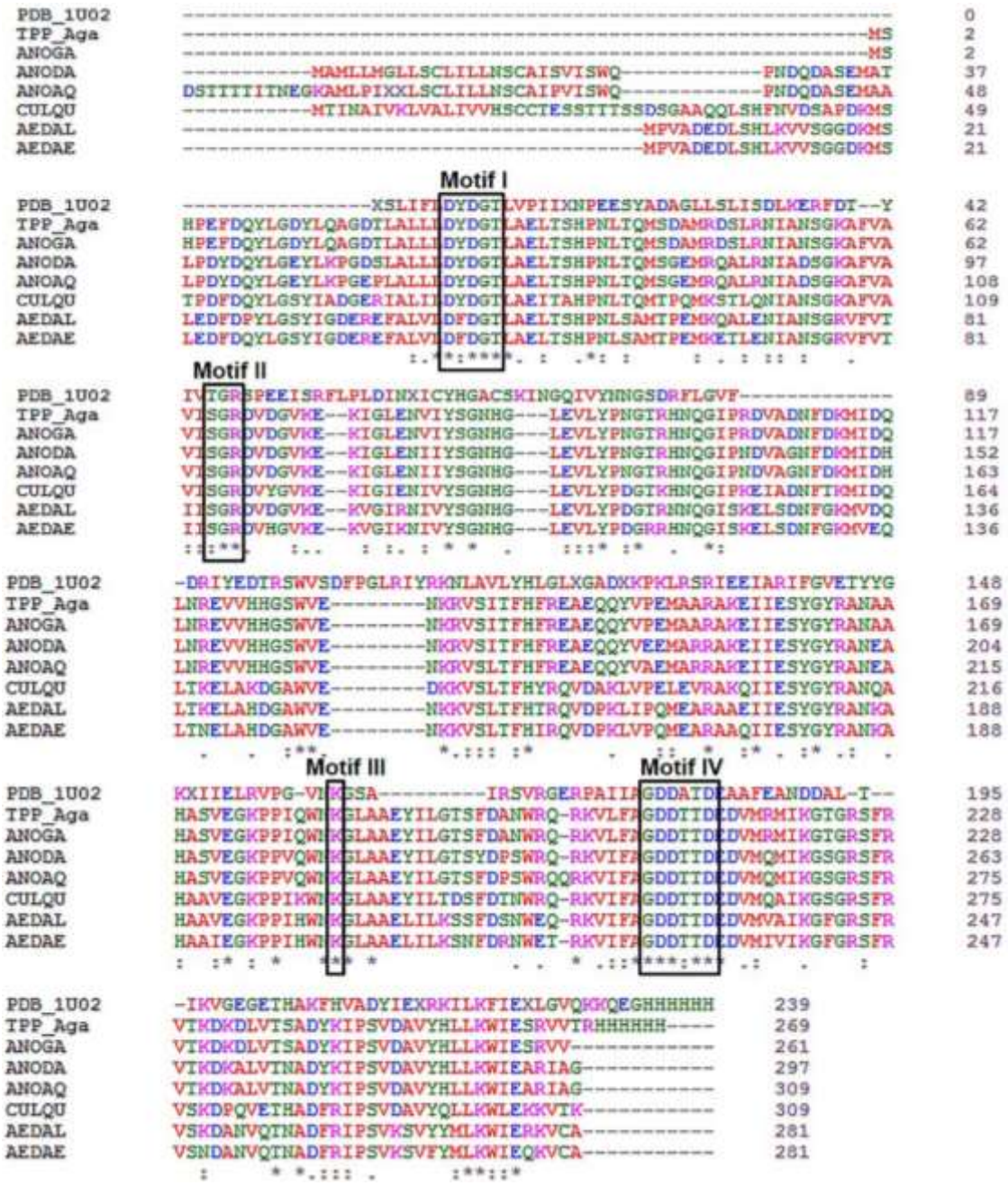

Figure 6. Alignment of the amino acid sequences of recombinant TPP with TPP sequences from different species of vector mosquitoes compared with that of the *T. acidophilum* template sequence. Identification of TPP sequences: PDB_1U02 (TPP from *T. acidophilum*), TPP_Aga (TPP from *A. gambiae* produced by *P. pastoris*), ANOGA (native TTP from *A. gambiae*), ANODA (from *Anopheles darlingi*), ANOAQ (from *Anopheles aquasalis*), CULQU (from *Culex quinquefasciatus*), AEDAL (from *Aedes albopictus*), AEDAE (from *Aedes aegypti*). Meaning of symbols under the sequences: ***** - identical regions; :- conserved regions; .- semiconserved regions. Regions with black rectangles indicate conserved motifs of the HAD superfamily: motif I - DXXX(T/V); motif II - (S/T)GX; motif III - K, motif IV - (G/S)(D/S)XXX(D/N).

motif II, where TPP from *T. acidophilum* has a threonine and insect TPP has a serine. In addition, the presence of a histidine tail at the TPP C-terminal portion does not appear to affect its biological function because it is distant from the functional motifs containing the catalytic residues.

One significant observation was the presence of a possible C2-type cap (Figure 7B). This cap is composed

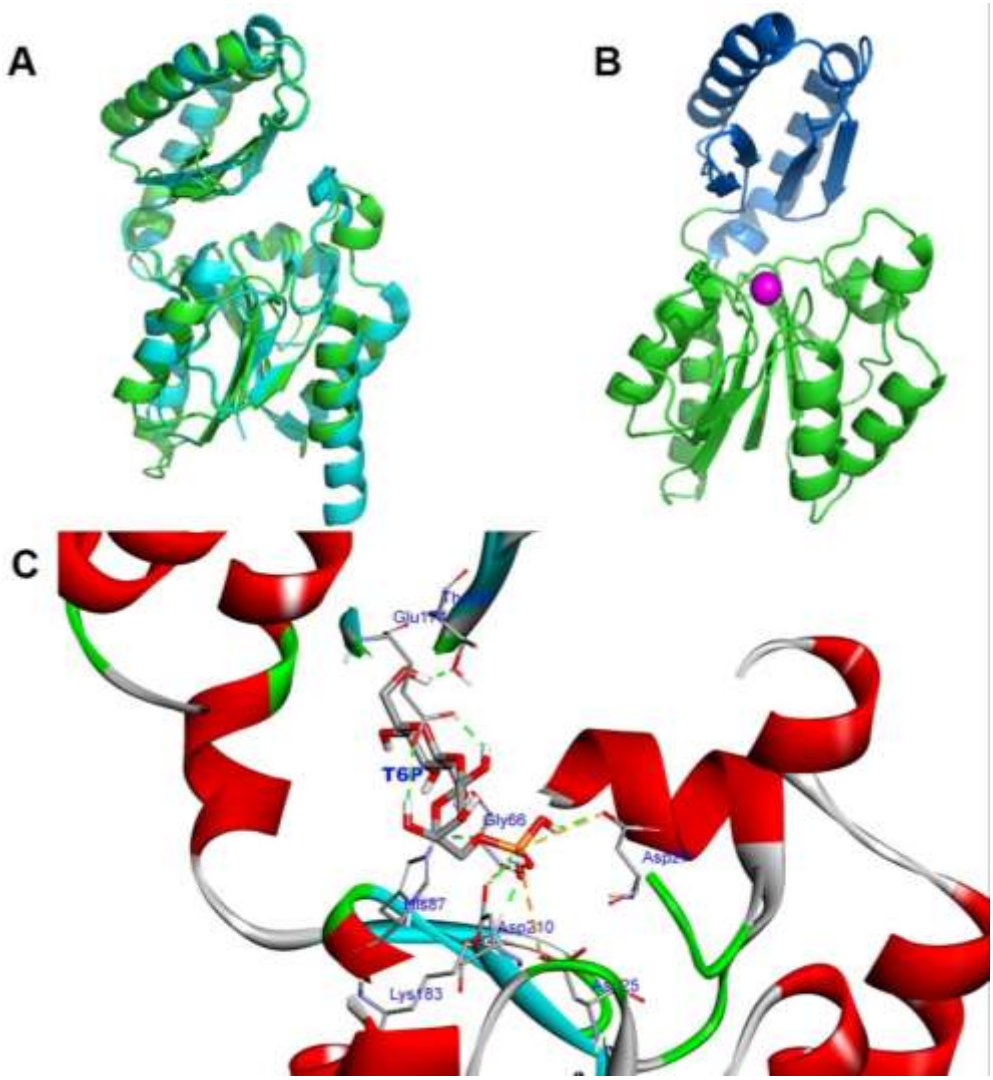

Figure 7. Structures of TPP from *A. gambiae* with their homologous model and simulation docking. **(A)** Ribbon diagram of the alignment of structural models for TPP from *A. gambiae* (green) and *T. acidophilum* (blue). The aligned structures were highly consistent between the two models. **(B)** Ribbon diagram of the tridimensional structure of TPP from *A. gambiae*, with the HAD central domain in green and the C2 cap domain in blue. Cofactor $Mg^{2+}$is indicated in magenta. **(C)** Simulation of T6P docking at the active site of TPP from *A. gambiae*. The T6P molecule represented in the line model is fit into the pocket formed between the central and cap domains. A detailed view of the active site shows the conserved residues involved in substrate binding. Green lines represent hydrogen bonds, and orange lines represent electrostatic interactions. Figure 7C was generated using the software AutoDock 4.2 and manipulated using the Discovery Studio 4.0 Visualizer.

of two consecutive units containing an α-helix followed by two β-strands (Figure 7B). The presence of this modification characterize a large clade within the HAD superfamily, which includes TPP and other phosphatases (Burroughs et al., 2006).

The tridimensional structure of recombinant TPP was obtained using the Phyre[2] web server. The stereochemical and general structural quality of the model generated by Phyre[2] was evaluated through an analysis of the Ramachandran plot using ProCheck, which showed that 92.9% of the residues estimated by the model were in the most favorable regions, 6.2% were in allowed regions, and 0.9% was in disallowed regions. Following validation, the tridimensional structure of TPP from *A. gambiae* was refined using the software ModRefiner and aligned with template 1U02 using the software PyMOL. The alignment of the two structural models is presented in Figure 7A.

The structural model exhibited 83% alignment coverage of the template 1U02, for which 223 amino acid residues were modeled with 100% confidence (probability that the *A. gambiae* sequences and template were homologues). The constructed model was based on an alignment generated using the Hidden Markov Model (HMM) maximum discrimination method (Eddy et al., 1995). The alignment quality was evaluated using the template modeling score, or TM-score (Zhang and Skolnick, 2005), which is a measure of structural similarity between two protein structures. The alignment of the structural model generated for *A. gambiae* TPP with the template revealed that the two models were consistent and indicated high reliability of the predicted structure for *A. gambiae* TPP.

The caps often contribute with residues that are required for specificity or auxiliary catalytic functions, and they play a central role in reactions catalyzed by the majority of HAD enzymes (Kurihara et al., 1995; Olsen et al., 1988). During the catalytic cycle, the cap domain moves across the catalytic site through binding interactions with the substrate-leaving group and forms a highly encapsulated active site (Farelli et al., 2014; Miao et al., 2016; Liu et al., 2017b). TPP typically has C2-type cap domains (in blue in Figure 7B) that project outward from the central domain (in green in Figure 7B). This type of domain is typical of the trehalose phosphatase family (Burroughs et al., 2006).

T6P binding to the TPP active site was modeled using the software AutoDock with a PyRx graphic interface (Figure 7C). The results were analyzed following protein visualization using Discovery Studio. Relevant information was obtained for the orientation of the enzyme-substrate binding interactions using AutoDock. For the docking simulations, the best binding position was the most energetically favorable one. The lowest binding free energy (ΔG_b) found for complex TPP-T6P was -4.30 kcal/mol, and the substrate molecule was stretched and interacted with the catalytic residues of the active site as expected. Therefore, the substrate fit well into the catalytic pocket formed at the interface of the central and cap domains near the active site residues. The substrate specificity for phosphatase is known to be dictated by the interface surface between these two domains (Farelli et al., 2014; Rao et al., 2006).

The interactions of T6P with TPP residues are presented in Figure 7C. The T6P sugar group is located at the cleft opening between the two domains, and it establishes hydrogen bonds with residues His87, Thr137 and Glu174, indicating that these residues may be involved in T6P binding. However, the phosphate group is located inside of the enzyme active site, where it can interact with catalytic residues Asp25, Asp27, Gly66, Lys183 and Asp210. The results obtained using AutoDock are consistent with the description of the main residues involved in interactions with the substrate. The interaction of the phosphate group with the catalytic residues of the active site is observed through hydrogen bonds, which are represented as dashed green lines, between the substrate oxygen (red tips of the phosphate group) and residues belonging to the HAD superfamily conserved motifs Asp27 (motif I), Gly66 (motif II), Lys183 (motif III) and Asp210 (motif IV). Electrostatic interactions (dashed orange lines) may occur between the substrate phosphorus and Asp residue (Asp25, Asp27 and Asp210) oxygen. The structure of the complex formed between the ligand and macromolecule supplies important information on the interactions that may occur at the active site, including hydrophobic and electrostatic interactions, hydrogen bonds and others. Figure 7C presents a detailed view of the catalytic site of TPP from *A. gambiae* and shows the evolutionary conservation of its component amino acids compared with the site of TPP from *T. acidophilum*. In addition, the similarity in spatial organization of the residues between the two enzymes was evident, indicating an affinity for the same type of substrate. Consequently, this pattern of T6P binding specificity is observed in other organisms that exhibit the trehalose biosynthesis pathway.

Because trehalose is an important carbohydrate for the metabolism of insects, studies have focused on the enzymes of its biosynthetic pathway to determine potential inhibitors. Kern et al. (2012) suggested that interference in trehalose synthesis could lead to the discovery of new insecticide action mechanisms because TPS (the other enzyme involved in trehalose synthesis) is also a potential drug target. The mechanism of inhibition of trehalose synthesis enzymes as a viable contribution to new insecticides is supported by previous studies (Chen et al., 2010; Tang et al., 2010). Trehalose was suggested as a likely source of energy for *Plasmodium* in *A. gambiae* mosquitoes after they feed on blood infected with these pathogens, since the parasite directly captures and metabolizes trehalose or is hydrolyzed to glucose and then captured by *Plasmodium* (Liu et al., 2013).

Recently, important drug target candidates were identified in the *B. malayi* nematode, that cause human filariasis, and TPP was one of the most important selected targets (Farelli et al., 2014). A vaccine is currently being developed using recombinant TPP from *B. malayi* with the goal of decreasing or preventing the development of the lymphatic filariasis parasite, and it has shown promising preliminary results (Kushwaha et al., 2013). Therefore, TPP is a key enzyme that plays a fundamental role in insect metabolism and physiology, and its absence in mammalian cells makes it a target for insecticide development.

Conclusions

The results of the present study show that the *P. pastoris* expression system provided efficient and stable source of TPP from *A. gambiae*. The purified recombinant enzyme presented good activity for its specific substrate.

Homology modeling confirmed that TPP is a monomer and has a similar structure to the enzyme's crystallography deposited at the data bank, conserving the catalytic motifs of the active site.

As the trehalose metabolism is essential for insects and in this work several structural and kinetic aspects were elucidated for the first time of one of the important enzymes involved in this metabolic pathway from *A. gambiae*, the results increase the knowledge of the biochemistry and biology of this mosquito, that is the main vector of malaria on the African continent.

Based on the present study, it will also be possible to screen for new inhibitors assaying directly drugs or extracts against TPP or using its 3D modeled structure to do virtual screening by docking, and this can be the basis for the development of new insecticides to control mosquito vectors of important diseases.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

ACKNOWLEDGEMENTS

The authors wish to thank Dr. Luís André Morais Mariúba from the Leonidas and Maria Deane Research Institute from the Oswaldo Cruz Foundation - Amazon (ILMD-Fiocruz/AM) for their help with the immunoassay tests. This study was financed by the National Council for Scientific and Technological Development (CNPq), Coordination for the Improvement of Higher Education Personnel (CAPES) and the Amazon Research Support Foundation (FAPEAM).

REFERENCES

Accelrys Software (2013). Discovery studio modeling environment. Release 4.0 Accelrys Software Inc. San Diego.

Baker AS, Ciocci MJ, Metcalf WW, Kim J, Babbitt PC, Wanner BL, Martin BM, Dunaway-Mariano D (1998). Insights into the mechanism of catalysis by the P-C bond-cleaving enzyme phosphonoacetaldehyde hydrolase derived from gene sequence analysis and mutagenesis. Biochemistry 37(26):9305-9315.

Behm CA (1997). The role of trehalose in the physiology of nematodes, Int. J. Parasitol. 27(2):215-229.

Bolton EE, Wang Y, Thiessen PA, Bryant SH (2008). PubChem: integrated platform of small molecules and biological activities. In. Wheeler R, Spellmeyer D, editors. Annual Reports in Computational Chemistry. American Chemical Society: Washington. Pp. 217-241.

Burroughs AM, Allen KN, Dunaway-Mariano D, Aravind L (2006). Evolutionary genomics of the HAD superfamily: understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes. J. Mol. Biol. 361(5):1003-1034.

Cabib E, Leloir LF (1958). The biosynthesis of trehalose phosphate. J. Biol. Chem. 231(1):259-275.

Campbell MK, Farrell SO (2007). Bioquímica. Thomson Editora.

Cereghino JL, Cregg JM (2000). Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. FEMS Microbiol. Rev. 24(1):45-66.

Chen J, Zhang D, Yao Q, Zhang J, Dong X, Tian H, Chen J, Zhang W (2010). Feeding-based RNA interference of a trehalose phosphate synthase gene in the brown planthopper, *Nilaparvata lugens*. Insect Mol. Biol. 19(6):777-786.

Collet JF, Gerin I, Rider MH, Veiga-da-Cunha M, van Schaftingen E (1997). Human L-3-phosphoserine phosphatase: sequence, expression and evidence for a phosphoenzyme intermediate. FEBS Lett. 408(3):281-284.

Collet JF, Stroobant V, Pirard M, Delpierre G, van Schaftingen E (1998). A new class of phosphotransferases phosphorylated on an aspartate residue in an amino-terminal DXDX(T/V) motif. J. Biol. Chem. 273(23):14107-14112.

Cregg JM, Tolstorukov I, Kusari A, Sunga J, Madden K, Chappell T (2009). Expression in the yeast *Pichia pastoris*. Methods Enzymol. 463:169-188.

DeLano WL (2006). An Executable PyMOL Incentive Product Based on PyMOL v0.99. Delano Scientific LLC: South San Francisco.

Eastmond PJ, van Dijken AJ, Spielman M, Kerr A, Tissier AF, Dickinson HG, Jones JD, Smeekens SC, Graham IA (2002). Trehalose-6-phosphate synthase 1, which catalyses the first step in trehalose synthesis, is essential for *Arabidopsis* embryo maturation. Plant J. 29(2):225-235.

Edavana VK, Pastuszak I, Carroll JD, Thampi P, Abraham EC, Elbein AD (2004). Cloning and expression of the trehalose-phosphate phosphatase of *Mycobacterium tuberculosis*: comparison to the enzyme from *Mycobacterium smegmatis*. Arch. Biochem. Biophys. 426(2):250-257.

Eddy SR, Mitchison G, Durbin R (1995). Maximum discrimination hidden Markov models of sequence consensus. J. Comput. Biol. 2(1):9-23.

Elbein AD (2009). Cytoplasmic carbohydrate molecules: trehalose and glycogen. In. Moran AP, Holst O, Brennan P, von Itzstein M, editors. Microbial Glycobiology: Structures, Relevance and Applications. Academic Press: London. P. 1020.

Elbein AD, Pan YT, Pastuszak I, Carroll D (2003). New insights on trehalose: a multifunctional molecule. Glycobiology 13(4):17R-27R.

Farelli JD, Galvin BD, Li Z, Liu C, Aono M, Garland M, Hallett OE, Causey TB, Ali-Reynolds A, Saltzberg DJ, Carlow CK, Dunaway-Mariano D, Allen KN (2014). Structure of the trehalose-6-phosphate phosphatase from *Brugia malayi* reveals key design principles for anthelmintic drugs. PLOS Pathog. 10(7):e1004245.

Forattini OP (2002). Culicidologia médica: Identificação, Biologia e Epidemiologia. 2nd vol. USP: São Paulo.

Friedman S (1966). Trehalose-6-phosphate phosphatase from insects. Methods Enzymol. 8:372-374.

Friedman S, Hsueh T (1979). Insect trehalose-6-phosphatase: the unactivated type, as illustrated in *Periplaneta americana*, and a survey of the ordinal distribution of the two presently known types. Comp. Biochem. Physiol. 64B(4):339-344.

Garcia LS (2010). Malaria. Clin. Lab. Med. 30(1):93-129.

Gravel P (2002). Protein blotting by the semidry method. In. Walker JM, editor. The Protein Protocols Handbook. Humana Press: Totowa. Pp. 321-334.

Holleman AF, Wiberg E (2001). Inorganic Chemistry. Academic Press: Zurich, Switzerland.

Karthik L, Malathy P, Trinitta A, Gunasekaran K (2011). Molecular docking with trehalose-6 phosphate phosphatase: A potential drug target of filarial parasite "*Brugia malayi*". Lett. Drug Des. Discov. 8(4):363-370.

Kelley LA, Sternberg MJ (2009). Protein structure prediction on the web: a case study using the Phyre server. Nat. Protoc. 4(3):363-371.

Kern C, Wolf C, Bender F, Berger M, Noack S, Schmalz S, Ilq T (2012). Trehalose-6-phosphate synthase from the cat flea *Ctenocephalides felis* and *Drosophila melanogaster*: gene identification, cloning, heterologous functional expression and identification of inhibitors by high throughput screening. Insect Mol. Biol. 21(4):456-471.

Klowden MJ (2013). Physiological Systems in Insects. 3rd ed. Academy Press: Moscow.

Klutts S, Pastuszak I, Edavana VK, Thampi P, Pan YT, Abraham EC, Carroll JD, Elbein AD (2003). Purification, cloning, expression, and properties of mycobacterial trehalose-phosphate phosphatase. J. Biol. Chem. 278(4):2093-2100.

Koonin EV, Tatusov RL (1994). Computer analysis of bacterial haloacid dehalogenases defines a large superfamily of hydrolases with diverse specificity. Application of an iterative approach to database search. J. Mol. Biol. 244(1):125-132.

Kormish JD, McGhee JD (2005). The *C. elegans* lethal gut-obstructed *gob-1* gene is trehalose-6-phosphate phosphatase. Dev. Biol. 287(1):35-47.

Kurihara T, Liu JQ, Nardi-Dei V, Koshikawa H, Esaki N, Soda K (1995). Comprehensive site-directed mutagenesis of L-2-halo acid dehalogenase to probe catalytic amino acid residues. J. Biochem. 117(6):1317-1322.

Kushwaha S, Singh PK, Rana AK, Misra-Bhattacharya S (2011).Cloning, expression, purification and kinetics of trehalose-6-phosphate phosphatase of filarial parasite *Brugia malayi*. Acta Trop. 119(2-3):151-159.

Kushwaha S, Singh PK, Rana AK, Misra-Bhattacharya S (2013). Immunization of *Mastomys coucha* with *Brugia malayi* recombinant trehalose-6-phosphate phosphatase results in significant protection against homologous challenge infection. PLOS One 8(8):e72585.

Laemmli UK (1970).Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227(5259):680-685.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993). Procheck: A program to check the stereochemical quality for assessing the accuracy of protein structures. J. Appl. Crystallogr. 26(2):283-291.

Li YT, Zhang HH, Sheng HM, An LZ (2012). Cloning, expression and characterization of trehalose-6-phosphate phosphatase from a psychotropic bacterium, *Arthrobacter* strain A3. World J. Microbiol. Biotechnol. 28(8):2713-2721.

Liu C, Dunaway-Mariano D, Mariano PS (2017a). Rational design of first generation inhibitors for trehalose 6-phosphate phosphatases. Tetrahedron 73(10):1324-1330.

Liu C, Dunaway-Mariano D, Mariano PS (2017b). Rational design of reversible inhibitors for trehalose 6-phosphate phosphatases. Eur. J. Med. Chem. 128:274-286.

Liu K, Dong Y, Huang Y, Rasgon JL, Agre P (2013). Impact of trehalose transporter knockdown on *Anopheles gambiae* stress adaptation and susceptibility to *Plasmodium falciparum* infection. Proc. Natl. Acad. Sci. 110(43):17504-17509.

Mamedov TG, Suzuki K, Miura K, Kucho Ki K, Fukuzawa H (2001). Characteristics and sequence of phosphoglycolate phosphatase from a eukaryotic green alga *Chlamydomonas reinhardtii*. J. Biol. Chem. 276(49):45573-45579.

Matula M, Mitchell M, Elbein AD (1971). Partial purification and properties of a highly specific trehalose phosphate phosphatase from *Mycobacterium smegmatis*. J. Bacteriol. 107(1):217-222.

Miao Y, Tenor JL, Toffaletti DL, Washington EJ, Liu J, Shadrick WR, Schumacher MA, Lee RE, Perfect JR, Brennan RG (2016). Structures of trehalose-6-phosphate phosphatase from pathogenic fungi reveal the mechanisms of substrate recognition and catalysis. Proc. Natl. Acad. Sci. USA. 113(26):7148-7153.

Morais MC, Zhang W, Baker AS, Zhang G, Dunaway-Mariano D, Allen KN (2000). The crystal structure of *Bacillus cereus* phosphonoacetaldehyde hydrolase: insight into catalysis of phosphorus bond cleavage and catalytic diversification within the HAD enzyme superfamily. Biochemistry 39(34):10385-10396.

Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J. Comput. Chem. 19(14):1639-1662.

Nobre A, Alarico S, Fernandes C, Empadinhas N, da Costa MS (2008). A unique combination of genetic systems for the synthesis of trehalose in *Rubrobacter xylanophilus*: properties of a rare actinobacterial TreT. J. Bacteriol. 190(24):7939-7946.

Oakley BR, Kirsch DR, Morris NR (1980). A simplified ultrasensitive silver stain for detecting proteins in polyacrylamide gels. Anal. Biochem. 105(1):361-363.

Olsen DB, Hepburn TW, Moos M, Mariano PS, Dunaway-Mariano D (1988). Investigation of the *Bacillus cereus* phosphonoacetaldehyde hydrolase. Evidence for a Schiff base mechanism and sequence analysis of an active-site peptide containing the catalytic lysine residue. Biochemistry 27(6):2229-2234.

Peeraer Y, Rabijns A, Collet JF, van Schaftingen E, de Ranter C (2004).How calcium inhibits the magnesium-dependent enzyme human phosphoserine phosphatase. Eur. J. Biochem. 271(16):3421-3427.

Rao KN, Kumaran D, Seetharaman J, Bonanno JB, Burley SK, Swaminathan S (2006). Crystal structure of trehalose-6-phosphate phosphatase-related protein: biochemical and biological implications. Protein Sci. 15(7): 1735-1744.

Segel IH (1979). Bioquímica: teoria e problemas. Livros Técnicos e Científicos Editora: Rio de Janeiro.

Seo HS, Koo YJ, Lim JY, Song JT, Kim CH, Kim JK, Lee JS, Choi YD (2000). Characterization of a bifunctional enzyme fusion of trehalose-6-phosphate synthetase and trehalose-6-phosphate phosphatase of *Escherichia coli*. Appl. Environ. Microbiol. 66(6):2484-2490.

Shan S, Min H, Liu T, Jiang D, Rao Z (2016). Structural insight into dephosphorylation by trehalose 6-phosphate phosphatase (OtsB2) from *Mycobacterium tuberculosis*. FASEB J. 30(12):3989-3996.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7(1):539.

Silva Z, Alarico S, da Costa MS (2005). Trehalose biosynthesis in *Thermus thermophilus* RQ-1: biochemical properties of the trehalose-6-phosphate synthase and trehalose-6-phosphate phosphatase. Extremophiles 9(1):29-36.

Solis FJ, Wets RJ (1981). Minimization by random search techniques. Math. Oper. Res. 6(1):19-30.

Tang B, Chen J, Yao Q, Pan Z, Xu W, Wang S, Zhang W (2010).Characterization of a trehalose-6-phosphate synthase gene from *Spodoptera exigua* and its function identification through RNA interference. J. Insect Physiol. 56(7):813-821.

Tang B, Yang M, Shen Q, Xu Y, Wang H, Wang S (2017). Suppressing the activity of trehalase with validamycin disrupts the trehalose and chitin biosynthesis pathways in the rice brown planthopper, *Nilaparvata lugens*. Pest. Biochem. Physiol. 137:81-90.

Thompson SN (2003). Trehalose - the insect “blood” sugar. Adv. Insect Physiol. 31:205-285.

Vaidyanathan R, Estep A, Becnel J, Moore J, Talcott C (2015). Silencing trehalose-6-phosphate synthase incapacitates adult mosquitoes by interfering with the biosynthetic pathway for flight fuel. Mosquito and Fly Research. Agricultural Research Service: USDA.

Vieira R (2003). Fundamentos de bioquímica. UFPA: Belém.

Wolf LK (2009). PyRx. C&EN 87:31.

WHO (World Health Organization)(2016). World Malaria Report (2016). [http://www.who.int/malaria/publications/world-malaria-report-2016/report/en/].

Xu D, Zhang Y (2011). Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. Biophys. J. 101(10):2525-2534.

Zhang Y, Skolnick J (2005). TM-align: a protein structure alignment algorithm based on the TM-score. Nucleic Acids Res. 33(7):2302-2309.

academicJournals

Vol. 16(25), pp. 1379-1385, 21 June, 2017
DOI: 10.5897/AJB2016.15824
Article Number: A0BA7CB64864
ISSN 1684-5315

http://www.academicjournals.org/AJB

African Journal of Biotechnology

Full Length Research Paper

Leaf conditioning of Brazilian Cerrado species for DNA microextraction

Luiza Gabriela Fulgêncio de Lima[1], Caio César de Oliveira Pereira[2], Camila Gracyelle de Carvalho Lemes[1], Anderson Rodrigo da Silva[2], Juliana Oliveira da Silva[2], Dieferson da Costa Estrela[1], Guilherme Malafaia[1] and Ivandilson Pessoa Pinto de Menezes[1*]

[1]Department of Biology, IF Goiano, Urutaí, Goiás, Brazil.
[2]Department of Agronomy, IF Goiano, Urutaí, Goiás, Brazil.

Received 6 December, 2016; Accepted 9 June, 2017

Proper conditioning of leaf tissues in collection expeditions can affect the quantity and quality of DNA in the extraction process. The aim of this work was to define a method of preserving foliar tissue suitable for obtaining DNA from Brazilian Cerrado trees. Young leaves of species (Mangaba and Baru) were collected and conditioned in five different treatments during a period of six days. Genomic DNA was obtained using two alternative versions of the cetyltrimethyl ammonium bromide (CTAB) protocol. For Mangaba, no statistical differences were verified between means of DNA values obtained with diversity arrays technology (DArT) (55 ng/µL) and CTAB (48 ng/µL) methods. It was found that the amounts of DNA obtained with the methods used differed according with the conditioning type and time ($F_{20,60} = 1.98$; $p = 0.022$). For Baru, the mean of DNA extracted was significantly higher ($F_{1,60} = 42.81$; $p < 0.01$) from the CTAB method (80 ng/µL). A significant difference ($p < 0.05$) was also observed between DNA means of conditioning types ($F_{4,60} = 1.1$, $p = 2 \times 10^{-4}$), without this being detected over time. Any preservation method tested is indicated for the selection of Mangaba and Baru foliar tissue conditioning for DNA extraction in a short period (up to six days).

Key words: Conservation, germplasm, native tree, DNA purification.

INTRODUCTION

The Cerrado is considered one of the world's biodiversity hotspots, housing 1/3 of the Brazilian biota and 5% of the world's flora and fauna (Eiten, 1994; Sloan et al., 2014). In the area, it is the second largest vegetation cover in Brazil and South America (Mendonça et al., 1998; Sano et al., 2010) with a high impact on food security (FAO, 2015). This large territory, according to the United States, Department of Agriculture, allowed Brazil to achieve high levels of agricultural production, making it the largest exporter of beef, chicken, sugar cane and ethanol in the world and second to soybean, making clear the functional assets that Cerrado offers us. The model of agricultural expansion used is leading to a progressive depletion of its natural resources (Machado et al., 2004; Silva et al.,

*Corresponding author. E-mail: ivan.menezes@ifgoiano.edu.br.

Table 1. Description of five treatments of Mangabeira and Baruzeiro leaf tissue in the collection for DNA extraction.

Treatment	Description of foliar conditioning	Acronym
01	Leaf conditioned in falcon tube (50 ml) with 40 ml of TE (Tris-EDTA buffer) (1 M Tris HCl pH = 8.0, 0.5 M EDTA pH = 8.0), kept at ambient temperature and without light.	TEB
02	Leaf conditioned in falcon tube (50 ml) with 40 ml of TE buffer (1 M Tris HCl pH = 8.0, 0.5 M EDTA pH = 8.0), kept without light, collection in ice (I) and kept at -20°C until extraction;	TEBI
03	Leaf conditioned in Falcon tube (50 ml), kept without light, collection in ice (I) and kept at -20°C until extraction;	I
04	Leaf conditioned in Falcon tube (50 ml), with 10 ml of silica gel (Sg) and kept without light;	Sg
05	Leaf conditioned in Falcon tube (50 ml), without substance (WS) and kept without light;	WS

2006), ranking the Cerrado as one of the most endangered regions of the planet (Hoekstra et al., 2005).

About 80% of the biodiversity in this area has already undergone severe changes in its natural space, intensifying the constant threats of local extinctions and concerns about the maintenance and viability of the natural services offered (Françoso et al., 2015; WWF, 2014). Thus, it is considered important, the preservation of this heritage and the relevance of a sustained exploitation of its natural resources. For this, the use of the productive potential of its native fruit trees is recommended (Machado et al., 2004).

Until then, approximately 60 native fruit tree species are known and traditionally used by families living in the Cerrado (Franzon, 2009). According to the author, its use is still essentially extractive and often predatory, demonstrating the importance of its cultivation. In most cases, lack of knowledge of the distribution of genetic variability, propagation techniques and phenology of the species makes it difficult to grow commercially. This fact makes it necessary that basic knowledge about the biological diversity of the Cerrado be consolidated and disclosed.

The establishment of conservation plans and adequate management is essential. Studies on the genetic makeup of native populations will assist and support the design of sampling and use strategies. DNA markers have been used successfully for this purpose (Desalle and Amato, 2004). The good quality of extracted genomic DNA is a major step in obtaining a DNA marker via Polymerase Chain Reaction (PCR).

For purification of plants DNA, different DNA extraction protocols have been used (Edwards et al., 1991; Doyle, 1990; Dellaporta et al., 1983). Certain protocols are not suitable for some species, which does not lead to obtaining a sufficient amount of DNA or, when obtained, presenting a very low purity, which may lead to problems and amplification failures (Romano and Brasileiro, 1999), since different species have specific biochemical behaviors, such as the release of free and secondary radicals, which reduce the quality and quantity of DNA extracted (Cavallari et al., 2014).

The adequacy of extraction techniques is often carried out for several species of plants (Edge-Garza et al., 2014; Lade et al., 2014; Mogni et al., 2016), but few or no reference has been made to the proper way of conditioning the foliar tissues, obtained in collection expeditions. The way to preserve the tissue for DNA micro extraction is as important as the choice of the DNA extraction protocol so that, optimizations are necessary to facilitate the operation and ensure a better quality of extracted DNA (Tamari and Hinkley, 2016).

The literature is lacking in studies related to the native species of the Cerrado, so the determination of an ideal way to conserve the tissues, as well as a specific extraction protocol for such species is important, to promote the obtaining of genetic material for molecular studies. The objective of this work was to determine a suitable conditioning method for the preservation of foliar tissues samples, collected from the Cerrado for purification of genomic DNA using two alternative versions of cetyltrimethyl ammonium bromide (CTAB) protocol.

MATERIALS AND METHODS

Plant and conditioning

The Mangabeira and Baruzeiro trees were the species chosen for study. The reason for this choice was the contrast in the leaves of each plant, as to the external morphological aspects.

The leaf of the Mangabeira is thick, hairy and chartaceous, and the presence of latex, whereas the one of the Baruzeiro is thin, leathery and less hydrated. Six young and whole leaves were collected from a single plant using 50 mL falcon tubes and conditioned separately in each treatment for six days after collection (Table 1). Five treatments were defined based on the methods of preservation of the samples of leaves used for collection expeditions until the moment of extraction, such as: 1) TE buffer; 2) TE buffer with ice and maintained at -20°C (TEBI); 3) with ice alone and maintained at -20°C (I); 4) only with silica gel; 5) without substance (WS). The accessions belong to the Germplasm

Bank *in vivo* of the Goiás State University Campus Ipameri and Instituto Federal Goiano Campus Urutaí.

DNA extraction and quantification

About 1 g of tissue corresponding to the leaf limb was used, which was macerated in a porcelain mortar, through mechanical maceration with the aid of liquid nitrogen. Part of the macerate was placed in a 2.0 ml plastic tube, occupying about ¼ of its volume.

The DNA was extracted separately using two DNA purification methods, one based on the protocol of 2% CTAB (Ferreira and Grattapaglia, 1996), often referenced in the literature; and a less common one called plant DNA extraction protocol for (diversity arrays technology) DArT, an alternative version of the 2% CTAB, recommended by Diversity Arrays Technology PTY LTD. The difference between the protocols is the composition of the extraction buffer used, and the 2% CTAB protocol employs the use of polyvinylpyrrolidone (PVP) and β-mercaptoethanol, whereas the working buffer presented by DArT employs the use of sorbitol in its composition.

The obtained genomic DNA was diluted in 100 µL Milli-Q autoclaved water. Quantification was performed at 0.8% (w/v) agarose gels submitted to electrophoresis. 2.0 µl aliquots of DNA from each obtained sample were applied to the gel wells beside a series of known concentrations of λ phage DNA (50, 100 and 200 ng). Sample concentrations were estimated by visual comparison of the fluorescence intensity of the λ phage DNA bands. Gels were visualized after staining with ethidium bromide (0.5%/ml) in TBE buffer solution (1X) for 10 min.

Data analysis

The quantification data of extracted DNA were tabulated and systematized in Excel and later submitted to the analysis of variance (ANOVA) according to the three-way ANOVA model, considering the extraction method, conditioning type and time factors, as well as the interactions between them.

The first two factors were studied by applying Fisher's least significant difference (LSD) test. The time factor was studied by adjusting polynomial models of first and second-degree, using as criterion the F-test and the coefficient of determination. Normality and residual homoscedasticity were previously checked. All inferential procedures were performed at the 10% level of significance. The analyses were performed using R software, version 3.2.1.

RESULTS AND DISCUSSION

Visual aspects of foliar tissues

Comparing the Mangaba foliar tissues submitted to the five methods of preservation, through the time intervals between the extracted DNA, visual changes were verified in the conditioned samples. However, it was not possible to establish a pattern that could be related to treatment or time in this short period of test.

So, for this short period, test measurements of the free radicals and oxidants would be necessary. Some samples showed color-related changes, presenting a green-yellow color with some punctual red spots, which were more evident in the samples of the fifth day. On the other days, conditioning was similar to those when they were collected. However, tissues treated with silica gel showed gradual dehydration over the period in which the extractions were performed. Total visual drying occurred in approximately 48 h (Figure 1).

Regarding the conditioning of the Baru foliar tissues, no visual changes in coloration were observed, nor were they stained with time on the samples kept in TEBI, TEB, I and WS, indicating an adequate state of conservation. The result is different from that observed in the conditioning of the Mangabeira leaf samples. On silica gel, the tissue dehydration process was also verified, however, faster than in Mangaba tree (Figure 2). The total visual drying was observed in approximately 14 h, demonstrating that it occurred more satisfactorily in Baru than in Mangabeira. Once the foliar samples are necessary more than 12 h after collection for drying, the chance of degradation of its DNA is increased, due to the accumulation of free radicals and oxidants (Sytsma et al., 1993).

The absence of visual morphological changes observed in the foliar tissues in Baru shows resistance to senescence that must be related to its thin, thickness and with the low water content, which may hinder the rapid accumulation of free radicals and oxidants after collection. This observation indicates that, there will hardly be any problem in conditioning the samples for a short period or the equivalent of one week for any of the methods chosen. Different from what has been observed for leaf tissue of Mangabeira, is thicker and juicier, indicating that care should be taken in the choice of conditioning since we can detect signs of alteration of the color of the limbus. This yellow-green is related to the early process of foliar senescence, due to the accumulation of ethylene and cytokinin reduction (Soares, 2008), which must be triggered by leaf removal. Based on these results, it was not possible to determine for sure which method of conditioning is suitable for Mangabeira, indicating that biochemical analysis also would be necessary.

Characteristics observed during the DNA extraction procedure

In the mechanical maceration stage of the tissues, the importance of the choice of young leaves for Mangabeira to avoid resistance in the grinding process was verified, due to the high density of secondary veins. Another highlight was the rapid darkening of the sample after maceration, a fact that is related to the oxidation process. This chemical reaction is damaging to DNA, since it leads to the release of free radicals, which causes the fragmentation of nucleic acids (Rajan et al., 2014).

These aspects were not observed in Baru DNA extractions from young leaf tissue. This result indicates, mainly for Mangabeira, the necessity of using DNA extraction protocols with optimized anti-oxidant concentrations, to decrease its degradation (Silva, 2010). In this context, it is necessary to test several experimental scenarios considering variations in the composition of the

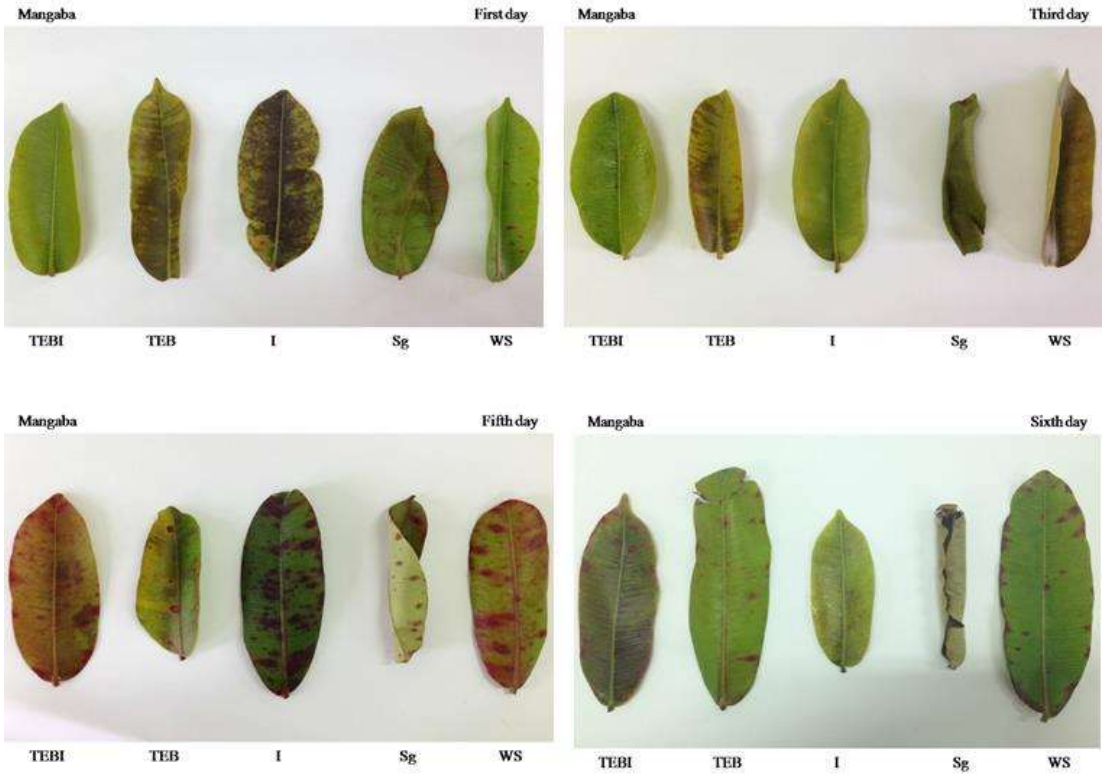

Figure 1. Visual aspects of Mangaba foliar tissues during the 1st, 3rd, 5th and 6th days of conditioning, through the five different packages used.

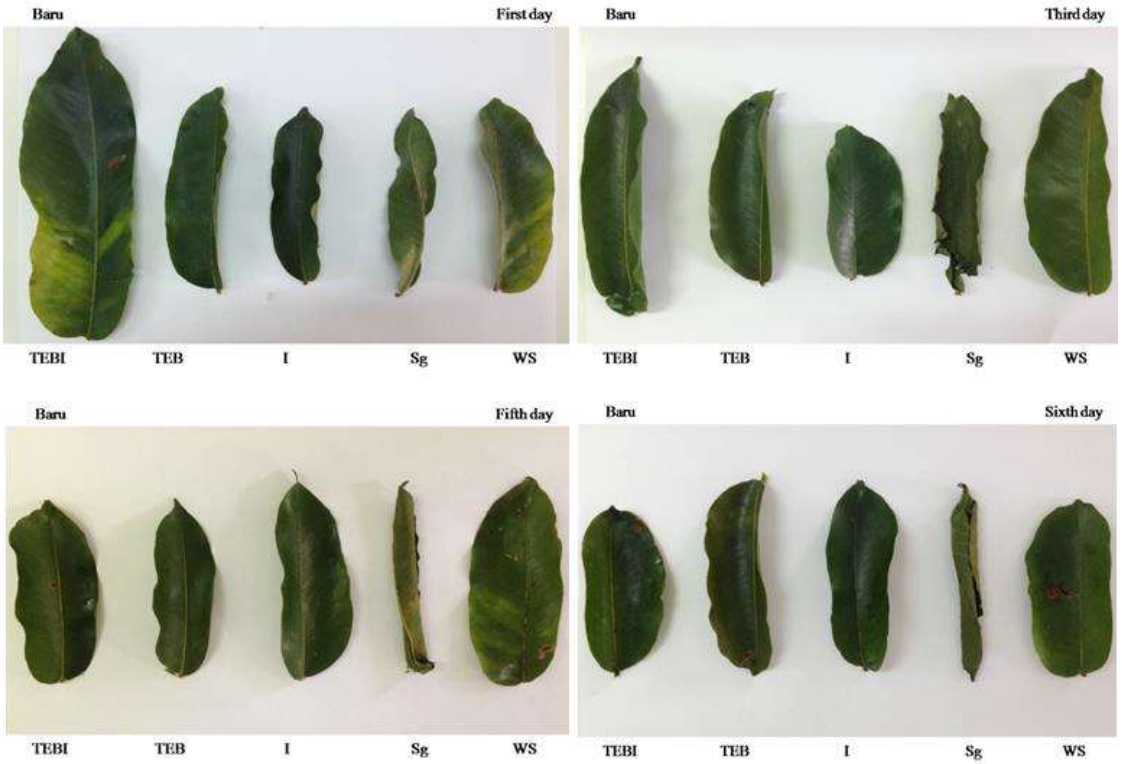

Figure 2. Visual aspects of Baru foliar tissues during the 1st, 3rd, 5th and 6th days of conditioning, through the five different packages used.

Table 2. Quantities of DNA extracted from Mangaba foliar tissue, considering the conditioning method over time according to the extraction protocol used. In the lines, upper case evidence the comparisons between the conditioning used, in the columns, lower case evidences the comparisons between the protocols used.

Time (h)	Extraction protocol	Conditioning				
		Sg	TEB	TEBI	I	WS
0	CTAB	$_{Aa}50$	$_{Aa}50$	$_{Aa}100$	$_{Aa}50$	$_{Aa}75$
	DArT	$_{Aa}15$	$_{Aa}15$	$_{Ab}40$	$_{Aa}15$	$_{Aa}50$
24	CTAB	$_{Aa}100$	$_{Aa}75$	$_{Aa}75$	$_{Aa}75$	$_{Aa}50$
	DArT	$_{ABb}50$	$_{Cb}0$	$_{Cb}0$	$_{Aa}75$	$_{BCa}15$
48	CTAB	$_{CDb}25$	$_{Db}0$	$_{Ab}100$	$_{ABa}75$	$_{BCb}50$
	DArT	$_{Ba}75$	$_{BCa}65$	$_{Aa}150$	$_{Cb}30$	$_{Aa}150$
72	CTAB	$_{Aa}75$	$_{Aa}75$	$_{Bb}30$	$_{Bb}30$	$_{Ba}0$
	DArT	$_{Ba}75$	$_{Ba}75$	$_{Ba}75$	$_{Aa}125$	$_{Ca}30$
96	CTAB	$_{Ba}0$	$_{Ab}50$	$_{Bb}0$	$_{Bb}0$	$_{Aa}50$
	DArT	$_{Ba}20$	$_{Aa}100$	$_{Ba}50$	$_{Ba}55$	$_{Ba}40$
120	CTAB	$_{Aa}100$	$_{Bb}50$	$_{Ca}0$	$_{BCa}30$	$_{Ca}0$
	DArT	$_{Aa}100$	$_{Aa}100$	$_{Bb}22.5$	$_{Ba}10$	$_{Ba}30$

extraction buffer as in the process (Lade et al., 2014).

Different from the other conditioning, the product of the extractions resulting from foliar tissues was maintained at -20°C. After the fourth day of conditioning, it showed a viscous appearance, mainly in the products from Mangaba tree foliar tissue samples. This viscous aspect is due to the presence of polysaccharides, an organic compound that hinders *in vitro* manipulations of DNA, such as PCR amplification and DNA cleavage due to inhibition of the action of DNA polymerase and restriction enzymes, respectively (Edwards et al., 1991; Fang et al., 1992; Sharma et al., 2002).

As for the pellet, the mass of extracted DNA was verified for the samples conditioned in silica gel. The smaller mass was about the other samples, both in Mangaba, as in Baru. This reduced size of the pellets certainly is associated with the lack of total maceration of the leaf tissue due to the stiffness of the dehydrated veins.

Estimation of quantities of obtained DNA

For the extractions carried out on Mangabeira leaf tissue, no statistically significant differences were observed for mean DNA values obtained using the DArT (55 ng/µL) and CTAB (48 ng/µL) methods. The amounts of DNA obtained showed wide amplitude with values 0 to 150 and 0 to 100 ng/µL of DNA, respectively. It was found that the amounts of DNA obtained from the extraction methods used differed according to the conditioning and time of preservation of the tissue ($F_{20,60}$ = 1.98, p = 0.022).

Considering this interaction frequently, larger amounts of DNA were observed among the products obtained from the DArT protocol than CTAB (Table 2). However, it was not possible to detect a significant association pattern (Figure 3). The lack of a pattern is related to the high variation of the data obtained in the micro window of pre-established observation time.

Meanwhile in Baru, the amount of DNA extracted on average was significantly higher ($F_{1, 60}$ = 42.81; p = 0), using the CTAB protocol (80 ng/µL) than DArT (46 ng/µL). There was also a significant difference in DNA means in the different conditions ($F_{4, 60}$ = 1.1; p = 2×10^{-4}). However, this difference was not detected over time. The amount of DNA obtained between the extraction methods used differed according to the conditioning and time, showing an influence of these factors on the amount of DNA obtained ($F_{20,60}$ = 1.69; p = 0.59) (Table 3). However, it was not possible to detect a clear association pattern between them in the observed observation time window (Figure 4).

Conclusion

For DNA extraction from Mangaba leaves, in short periods, the protocol DArT is the most effective. While for Baru, 2% CTAB protocol presents more efficient results,

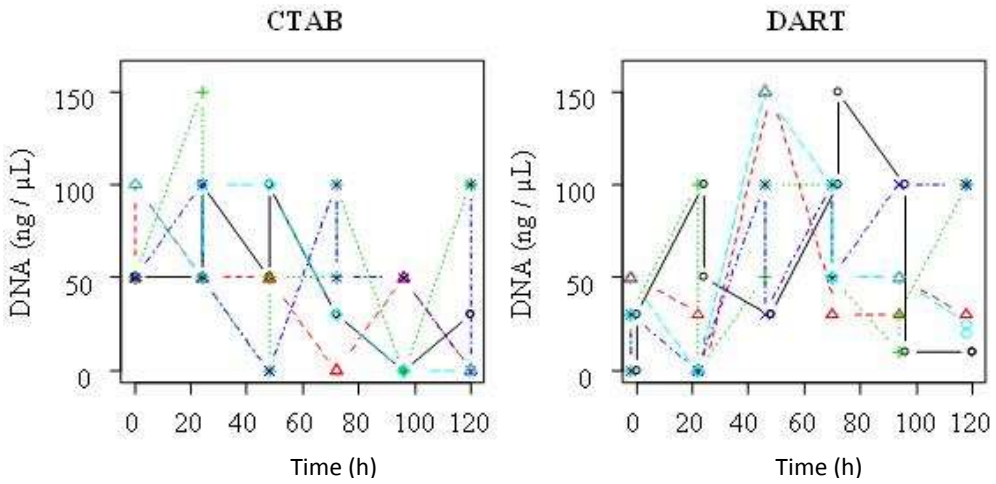

Figure 3. Graph of regression analysis of the amount of DNA separated by extraction method depending on conditioning process and preservation time for Mangaba. TEB, Dark blue line; TEBI, light blue line; I, Black line, Sg, green line; WS, red line.

Table 3. Quantities of DNA extracted from Baru foliar tissue, considering the conditioning method over time, according to the extraction protocol used. In the lines, upper case evidence the comparisons between the conditioning used, in the columns, lower case evidences the comparisons between the protocols used.

Time (h)	Extraction protocol	Conditioning				
		Sg	TEB	TEBI	I	WS
0	CTAB	$_{ABa}$75	$_{Aa}$100	$_{ABa}$75	$_{Ba}$50	$_{Cb}$0
	DArT	$_{Aa}$75	$_{Ab}$50	$_{Aa}$100	$_{Aa}$50	$_{Aa}$50
24	CTAB	$_{BCa}$40	$_{Ba}$50	$_{Aa}$100	$_{Aa}$100	$_{Cb}$0
	DArT	$_{Aa}$50	$_{Aa}$50	$_{Aa}$75	$_{Aa}$75	$_{Aa}$50
48	CTAB	$_{Aa}$90	$_{Aa}$50	$_{Aa}$100	$_{Aa}$45	$_{Aa}$100
	DArT	$_{Aa}$50	$_{Aa}$40	$_{Ab}$50	$_{Aa}$25	$_{Aa}$75
72	CTAB	$_{Aa}$100	$_{Aa}$100	$_{Aa}$100	$_{Aa}$100	$_{Aa}$100
	DArT	$_{Ab}$10	$_{Ab}$10	$_{Ab}$50	$_{Ab}$0	$_{Ab}$0
96	CTAB	$_{Aa}$100	$_{Aa}$100	$_{Aa}$150	$_{Aa}$80	$_{Aa}$100
	DArT	$_{ABb}$50	$_{Aa}$80	$_{BCb}$25	$_{Aa}$75	$_{Cb}$0
120	CTAB	$_{Aa}$100	$_{Ba}$50	$_{Aa}$125	$_{Ba}$50	$_{ABa}$80
	DArT	$_{BCb}$25	$_{Ba}$50	$_{Aa}$100	$_{Ba}$50	$_{Cb}$0

for extraction of genomic DNA.

Any preservation method tested is indicated for the selection of Mangaba and Baru foliar tissue conditioning method for DNA extraction in a short period (up to six days) of conservation. So, equally simple methods of sample preservation can be defined according to the availability of financial resources and instruments of the laboratory, where the procedures were performed.

CONFLICT OF INTERESTS

The authors declared no conflict of interest.

ACKNOWLEDGEMENTS

The authors thank the Instituto Federal Goiano for

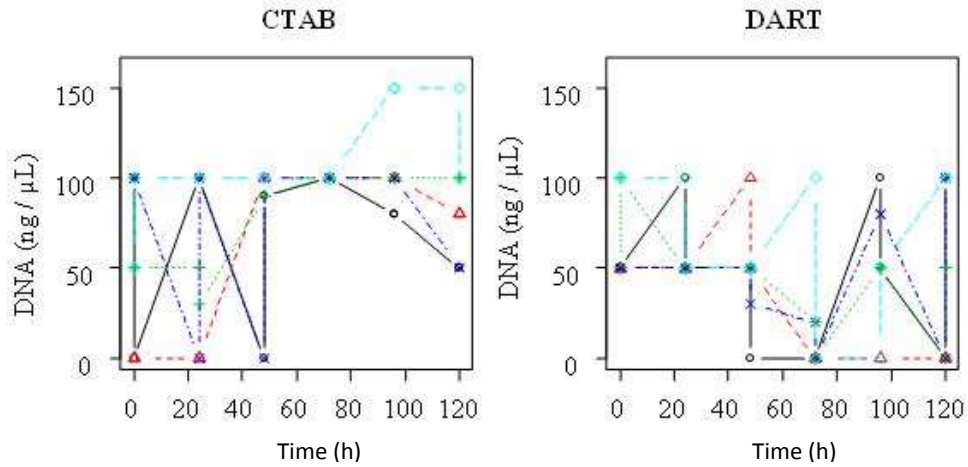

Figure 4. Graph of regression analysis of the amount of DNA separated by extraction method depending on conditioning process and preservation time for Baru. TEB, Dark blue line; TEBI, light blue line; I, Black line, Sg, green line; WS, red line.

REFERENCES

Cavallari MM, Siqueira MVBM, Val TM, Pavanelli JC, Monteiro M, Grando C, Pinheiro JB, Zucchi MI, Gimenes MA (2014). A modified acidic approach for DNA extraction from plant species containing high levels of secondary metabolites. Genet. Mol. Res. 13(3):6497-6502.

Dellaporta SL, Wood J, Hicks JB (1983). A plant DNA mini-preparation: version II. Plant Mol. Biol. Rep. 1(4):19-21.

Desalle R, Amato G (2004). The expansion of conservation genetics. Nat. Rev. Genet. 5(9):702-712.

Diversity Array Technology PTY LTD. Plant DNA extraction protocol for DArT. Available at: <http://www.diversityarrays.com>.

Doyle JJ (1990). Isolation of plant DNA from fresh tissue. Focus 12:13-15.

Edge-Garza DA., Rowland Jr TV, Haendiges S, Peace C (2014). A high-throughput and cost-efficient DNA extraction protocol for the tree fruit crops of apple, sweet cherry, and peach relying on silica beads during tissue sampling. Mol. Breed. 34(4):2225-2228.

Edwards K, Johnstone C, Thompson C (1991). A simple an drapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acid. Res. 19(6):1349.

Eiten GA (1994). A vegetação do Cerrado. In. Pinto MN (Ed). Cerrado: Caracterização, ocupação e perspectivas. 2.ed. Brasília: UNB, SEMATEC. pp. 9-65.

Fang G, Hammar S, Grumet R (1992). A quick and inexpensive method for removing polysaccharides from plant genomic DNA. Biotechniques 13(1):52-56.

FAO (2015). Agriculture conference: global food security and the role of sustainable fertilization. Available at http://www.fao.org/brasil/pt/.

Ferreira M, Grattapaglia D (1996). Introdução ao uso de marcadores moleculares em análise genética. 3 ed. Brasília: Embrapa. 220 p.

Françoso RD, Brandão R, Nogueira CC, Salmona YB, Machado RB, Collia GR (2015). Habitat loss and the effectiveness of protected areas in the Cerrado Biodiversity Hotspot. Nat. Conserv. 13(1):35-40.

Franzon RC (2009). Fruteiras nativas do Cerrado têm potencial para exploração. Planaltina, DF: Embrapa Cerrados. Available at https://www.embrapa.br/cerrados/publicacoes.

Hoekstra JM, Boucher TM, Ricketts TH, Roberts C (2005). Confronting a biome crisis: global disparities of habitat loss and protection. Ecol. Lett. 8(1):23-29.

Lade BD, Patil AS, Paikrao HM (2014). Efficient genomic DNA extraction protocol from medicinal rich Passiflora foetida containing high level of polysaccharide and polyphenol. SpringerPlus 3(1):457.

Machado RB, Ramos Neto MB, Pereira PGP, Caldas EF, Gonçalves DA, Santos NS, Tabor K, Steininger M (2004). Estimativas de perda da área do Cerrado brasileiro. Conservation International do Brasil, Brasília.

Mendonça RD, Felfili J, Walter B, Silva Júnior MD, Rezende A, Filgueiras T, Nogueira P, Sano S, Almeida SD (1998). Flora vascular do cerrado. In. Sano SM, Almeida SP. de. (Ed.). Cerrado: ambiente e flora. Planaltina, DF: Embrapa-CPAC Pp. 289-556.

Mogni VY, Kahan MA, Queiroz LP, Vesprini JL, Ortiz JPA, Prado DE (2016). Optimization of DNA extraction and PCR protocols for phylogenetic analysis in *Schinopsis* spp. and related Anacardiaceae. SpringerPlus 5(1):1-7.

Rajan I, Rabindran R, Jayasree PR, Kumar PRM (2014). Antioxidant potential and oxidative DNA damage preventive activity of unexplored endemic species of Curcuma. Indian J. Exp. Biol. 52:133-138.

Romano E, Brasileiro A (1999). Extração de DNA de plantas. Biotecnolog. Cienc. Desenvolv. 9:40-43.

Sano EE, Rosa R, Brito JL, Ferreira LG (2010). Land cover mapping of the tropical savanna region in Brazil. Environ. Monit. Assess. 166(1-4):113-124.

Sharma AD, Gill PK, Singh P (2002). DNA isolation from dry and fresh samples of polysaccharide-rich plants. Plant Mol. Biol. Rep. 20(4):415-415.

Silva J, Fariñas M, Felfili J, Klink C (2006). Spatial heterogeneity, land use and conservation in the cerrado region of Brazil. J. Biogeogr. 33(3):536-548.

Silva MN (2010). Extração de DNA genômico de tecidos foliares maduros de espécies nativas do cerrado. Rev. Árvore 34(6):973-978.

Sloan S, Jenkins CN, Joppa LN, Gaveau DLA, Laurance WF (2014). Remaining natural vegetation in the global biodiversity hotspots. Biol. Conserv. 177(2014):12-24.

Soares FP (2008). Influência das citocininas nos aspectos anatômicos, bioquímicos e fisiológicos do cultivo in vitro da mangabeira (*Hancornia speciosa* Gomes). Universidade Federal de Lavras, 2008.

Sytsma KJ, Givnish, TJ, Smith JF, Hahn WJ (1993). Collection and storage of land plant samples for macromolecular comparisons. MethodS Enzymol. 224:23-37.

Tamari F, Hinkley SC (2016). Extraction of DNA from Plant Tissue: Review and Protocols. In. Miéié M. (Ed). Sample Preparation Techniques for Soil, Plant, and Animal Samples. Springer Protocols Handbooks, pp. 245-263.

WWF (2014). Worldwide Fund for Nature. Living Planet report. Available at http://wwf.panda.org/.

academicJournals

Vol. 16(25), pp. 1386-1390, 21 June, 2017
DOI: 10.5897/AJB2017.15938
Article Number: 7CCF9F864866
ISSN 1684-5315

http://www.academicjournals.org/AJB

African Journal of Biotechnology

Full Length Research Paper

Effect of *Moringa oleifera* leaf extract on the haematological and serum biochemistry of rabbits reared in a semi-humid environment

Ojo, O. A.* and Adetoyi, S. A.

Department of Animal Production, Fisheries and Aquaculture, Kwara State University, Malete, Kwara State, Nigeria.

Received 12 February, 2017; Accepted 31 May, 2017

This research is aimed at improving the productive and pathological picture of rabbits reared in a semi-humid environment using *Moringa oleifera* leaf extract (MOLE). Twenty four mixed- breed rabbits, having an average weight of 700g, were used in this study. They were randomly divided into four equal treatments (6 rabbits each) and gavaged with various concentrations of MOLE. Treatment 1 (control) was given 0ml MOLE/kg body weight, treatment 2 (30 ml MOLE/kg body weight), treatment 3 (60 ml MOLE/kg body weight) and treatment 4 (90 ml MOLE/kg body weight). Results showed that *M. oleifera* leaf extract at all doses produced significant ($p<0.05$) changes in the blood levels of Packed Cell Volume (PCV), haemoglobin and white blood cell (WBC) count when compared to the control group. Rabbits given 30 of MOLE/kg body weight caused significantly ($p<0.05$) increased the recorded values of alkaline phosphatase (ALP). However, MOLE at tested dose of 90ml MOLE/kg body weight produced a significantly ($p<0.05$) lowered value in the serum level of alkaline phosphatase (ALP) with a non-significant change ($p>0.05$) observed in other serum parameters across the treatments. The total antioxidant capacity (TAC) value of rabbits increased consistently with increased MOLE concentration, while MDA values were not significantly influenced across the treatments. It can be concluded that (MOLE) can be used at 90ml MOLE/kg body weight to reduce lipid peroxidation and enhance oxidative status of rabbits in a semi-humid environment.

Key words: Gavaged, production, *Moringa oleifera*, rabbit, reproduction.

INTRODUCTION

Since ancient times, medicinal plants have been used for the control and treatment of human and livestock ailments (Ganesan and Bhatt, 2008). Such medicinal plants include, *Aloe ferox V* (Mwale et al., 2014), *Telfaria occidentalis* (Dada and Abiodun, 2014) and *Moringa oleifera* (Ojo et al., 2015). The effects of any feed ingredient including medicinal plants on the haematological factors of the livestock are of immense assistance in deciding whether or not such a feed ingredient will be safe as feedstuff (Mitruka and

*Corresponding author. E-mail: ykmelodya1@yahoo.com. Tel:+23408064221539.

Rawnsley, 1997). Oleforuh-Okoleh et al. (2015) stated that certain hematological factors such as packed cell volume, red blood cell, hemoglobin, etc., can be associated with certain production traits and serve as means of assessing clinical and nutritional health status of animals. It has also been documented that high packed cell volume (PCV) and high hemoglobin content (Hb) are associated with high feed conversion ratio (Mitruka and Rawnsley, 1997), while high percentage of white blood cells especially lymphocytes are associated with the ability of the animals to perform well under very stressful conditions. Medicinal plants, however, contain some toxins that have multi-system effects, such as acute kidney injury accompanied by hepatitis and colitis (Swanepoel et al., 2008). In some cases, however, medicinal plants do not have harmful effects on haematological and serum biochemical parameters (Jaouad et al., 2004; Oduola et al., 2007). According to Ewuola and Egbunike (2008), some medicinal plants are basically used as feed supplements, or for medicinal purposes thereby becoming involved in a cascade of physiological reactions, that may lead to alteration of haematological and serum biochemical parameters. This could result from the toxic substances that might be present in the plants in cases of lowering or elevating the haematological and biochemical values. It could also act as non-toxic invaluable compounds that maintain the values within the expected reference ranges for chickens (Simaraks et al., 2004). In view of this, the toxicological effects of *M. oleifera* on haematological and serum biochemical parameters and the effect on the oxidative status of rabbits naturally reared in a semi-humid environment were evaluated.

MATERIALS AND METHODS

This study was carried out at the Rabbitry section of the Teaching and Researc arm, College of Agriculture, KwaraStateUniversity, Ilorin, in Kwara State. The study lasted for 9 weeks.

Preparation of *M. oleifera* leaves extract (MOLE)

Fresh leaves of *M. oleifera* were collected early in the morning at Ita-alamu area of Ilorin, Kwara State. The *M. oleifera* leaves were manually removed from the stem, cleaned and made free of sand and other impurities using distilled water. The fresh leaves were blended into powdered using an electric kitchen blender. Finely pulverized *M. oleifera* leaves weighing 300 g was poured into a 2.5 L macerating flask and 1.5 L of distilled water added. The resulting mixture was thoroughly homogenized and sieved with a cheese cloth and then filtered using whatman filter paper (24 cm). Resulting filtrate was stored in the freezer (4 or -20°C) till use.

Experimental animals and management

Twenty four (24) grower rabbits of mixed breed rabbits aged9 months old, with average initial body weight ranging from 600 to 800 g were used for the experiment. The experimental animals were randomly assigned to four (4) groups comprising 6 animals per treatment in a completely randomized design. Each treatment was replicated thrice having 2 animals per replicate. The animals were allowed 10 days of acclimatization in the Teaching and Research Farm before data collection commenced. Prior to the data collection, the animals were given oxytetracycline (5%) intramuscularly twice, vitamin B complex intramuscularly and ivomec through subcutaneous route of administration. Treatment one (1) served as the control given 0 ml *M. oleifera* leaf extract /kg body weight of animal, while rabbits in treatments 2, 3 and 4 received the MOLE extract at 30 ml MOLE/kg body weight, 60 ml MOLE/kg body weight and 90 ml MOLE/kg body weight, respectively via gavaging method of administration. Feed and water were given *ad libitum*. The experimental design was completely randomised design (CRD).

Table 1. Chemical compositions of *Moringa oleifera* leaves.

Item	*M. oleifera* leaves (MOL)
Dry matter (%)	91.78
Moisture content	8.22
Crude protein (%)	28.43
Crude fat (%)	6.40
Crude fibre	9.15
Total ash (%)	9.09
Nitrogen free extract (%)	46.93%

Chemical analysis

Proximate analysis of experimental diets was carried out in a reputable chemical analysis labouratory using the method described by Association of Analytical Chemist (A.O.A.C, 1990).

Blood collection

2 ml blood sample was collected from the marginal ear vein of 12 rabbits which were comfortably restrained prior to the blood collection. The blood samples were collected in heparinised tubes, temporarily stored in crushed ice and transported to a reputable laboratory for analysis. Malondialdehide (MDA) an index of lipid peroxidation was determined using the method of Buege and Aust (1978). Total antioxidant capacity of serum sample was determined according to the method described by Baydar et al. (2007).

Statistical analysis

Data collected were subjected to one way analysis of variance (ANOVA) in a completely randomized design, while significant means were separated using Duncan's new multiple test of software (SAS, 2003).

RESULTS

Chemical analysis

The results of proximate analysis in Table 1 showed that Moringa leaves had an appreciable crude protein content (28.43%), crude fibre (9.15%), ash (9.09%), dry matter (91.78%), nitrogen free extract (NFE) (46.93%), but low

Table 2. Haematological parameters of rabbits fed MOLE.

Parameter	T_1 (0%)	T_2 (30%)	T_3 (60%)	T_4 (90%)
PCV (%)	32.00^{a}±2.52	25.50^{a}±1.50	22.33^{ab}±2.33	26.67^{ab}±2.40
HB (g/dl)	10.30^{a}±1.14	8.30^{ab}±0.30	7.33^{b}±0.58	8.66^{ab}±0.80
RBC ($\times10^{6}$)	5.16±0.58	4.12±0.06	3.56±0.34	4.25±0.47
WBC ($\times10^{3}$)	5.82^{b}±2.24	4.70^{bc}±4.00	4.25^{c}±6.64	8.05^{a}±3.3
Platelet ($\times10^{3/}$l)	88.67± 2.5	88.50±3.5	67.33±5.9	93.33±6.6
Lymphocytes (%)	62.00±4.35	66.00±0.00	68.33±2.73	67.00±1.15
Neutrophils (%)	33.67±4.33	29.50±0.50	28.67±2.3	29.67±0.6
Monocytes (%)	2.33±0.33	2.00±1.00	1.67±0.33	1.33±0.33
Eosinophils (%)	2.00±0.57	2.50±0.50	1.33±0.35	1.67±0.88

$^{a-b}$Means bearing different superscript in the same row differ significantly($P < 0.05$). T_1 = Control, T_2 = 30 ml of *M. oleifera* leaf extract (MOLE) /kg body weight, T_3 = 60 ml of *M. oleifera* leaf extract (MOLE) /kg body weight, T_4 = 90 ml of *M. oleifera* leaf extract (MOLE) /kg body weight; SEM, Standard error of means. NS = No significant different; S = Significant different; MDA = Lipid peroxidation; TAC = Total antioxidant capacity; K = potassium level.

content of ether extract (6.40%). Dry matter content of MOLE in this study was lower than the values reported by Mutayoba et al. (2011) who reported dry matter values of 93.7%. The crude protein (CP) value of MOLE obtained in this study was higher than the value reported by Olugbemi et al. (2010) which was 27.44%, although Mutayoba et al. (2011) recorded a higher (30.65%) crude protein value. The crude fat and ash values (6.40 and 9.09%) observed in this study were higher than the values 2.11 and 7.93% reported by Ogbe et al. (2011/2012). Crude fibre value of 9.15% reported in this study was higher than 5.43% which was reported in the study conducted by Sodamade et al. (2013). These differences in values of MOLE have been observed in the previous studies stated to be due to differences in the soil type, climatic conditions, stage of maturity and their genetic make-up.

Haematological analysis

The haematological parameters of rabbits given the varying concentrations of MOLE are presented in Table 2. Results showed that *M. oleifera* at all doses administered to the animals produced significant ($p>0.05$) change in the blood levels of packed cell volume, haemoglobin and total white blood cell count when compared to the control group. The lowest value of PCV was observed in rabbit group given 60ml MOLE/kg body weight (22.33±2.33%), while the highest value was recorded in the control group with 0ml MOLE/kg body weight (32.00 ±2.52%). 60ml MOLE/kg body weight significantly decreased the HB value compared with the control value (10.30±1.14 g/dl). MOLE administration at 90ml MOLE/kg body weight significantly lowered the WBC value at 60 ml MOLE/kg body weight (4.25×10^{3} ± 6.64) while the highest value was observed in rabbits given 90 ml MOLE/kg body weight (8.05×10^{3}±3.3). The RBC, platelet, lymphocytes, neutrophil, monocyte and eosinophil were not significantly affected among experimental groups.

Serum biochemical analysis

On the other hand, treatment of the animals with 30 ml of MOLE/kg body weight caused significantly ($p<0.05$) increased value of ALP. However, *M. oleifera* at tested dose of 90ml /kg body weight produced a significantly ($p<0.05$) lowered value in the serum level of ALP with a non-significant change ($p>0.05$) observed in the serum levels of albumin, globulin, albumin: globulin ratio (A: G ratio), total protein, total bilirubin, cholesterol, creatinine, blood urea nitrogen (BUN), aspartate transaminase, alanine amino transferase, calcium and sodium levels when compared to the control group (Table 3).

Table 4 shows the result of serum lipid peroxidation (MDA), total antioxidant capacity (TAC) of rabbit bucks fed *M. oleifera* leaf extract (MOLE). The values recorded for MDA ranged from 1.46 to 0.31 u/l with the highest value recorded in rabbits given no MOLE (control 0 ml MOLE /kg body weight) while the lowest value was observed in rabbits given 90 ml MOLE /kg body weight, although no significant ($p>0.05$) difference was observed in the MDA value across the treatments. The TAC recorded was significantly influenced by the MOLE administration. The TAC value of rabbits increased consistently with an increased MOLE concentration. The highest value (40.81 µg/ml) was recorded in rabbits given 90 ml MOLE/kg body weight, while the lowest value (28.32 µg/ml) was observed in the control group which received no MOLE.

DISCUSSION

The study investigated the effect of aqueous extract of *M.*

Table 3. The serum biochemical parameters of rabbits given oral administration of MOLE.

Parameter	T_1 (0%)	T_2 (30%)	T_3 (60%)	T_4 (90%)
Albumin(g/dl)	3.53±0.20	3.60±0.20	4.00±0.31	4.03±0.23
Globulin(g/dl)	4.50±0.35	4.75±0.15	5.17±0.27	5.00±0.10
A:G ratio	0.73±0.03	0.70±0.00	0.77±0.06	0.73±0.03
Total protein(g/dl)	8.03±0.54	8.35±0.35	9.17±0.33	9.03±0.23
Total Bilirubin(mg/dl)	0.30±0.11	0.15±0.05	0.23±0.06	0.20±0.00
Cholesterol(mg/dl)	20.67±6.66	19.50±1.56	25.67±8.98	23.0±5.50
Creatinine(mg/dl)	1.30±0.40	1.55±0.25	2.06±0.33	2.03±0.22
BUN(mg.dl)	14.10±0.40	15.45±0.95	15.90±0.73	15.93±0.33
AST(iu/l)	23.67±0.67	33.0±8.00	18.0±2.65	44.30±14.81
ALT(iu/l)	149.67±61.29	113.50±0.50	162.37±11.5	140.3±24.90
ALP(iu/l)	36.33^{ab}±0.88	39.0^{a}±1.00	33.0^{ab}±1.00	32.33^{b}±2.90
Ca(mg/dl)	13.07±1.34	12.70±0.90	13.07±1.59	13.67±1.47
Na(mg)	119.0±4.35	112.50±2.50	116.00±3.05	118.00±4.60

$^{a-b}$Means bearing different superscript in the same row differ significantly($P < 0.05$); T_1 = Control, T_2 = 30 ml of *M. oleifera* leaf extract (MOLE) /kg body weight, T_3 = 60 ml of *M. oleifera* leaf extract (MOLE)/kg body weight, T_4 = 90 ml of *M. oleifera* leaf extract (MOLE) /kg body weight; SEM, Standard error of means. NS = No significant different; S = Significant different; MDA = Lipid peroxidation; TAC = Total antioxidant capacity; K = potassium level.

Table 4. Lipid peroxidation, total antioxidant capacity of rabbit bucks fed MOLE.

Parameter	T_1	T_2	T_3	T_4	Std. Mean Error	P-value
MDA (u/l)	1.46	0.37	0.59	0.31	0.21^{NS}	0.126
TAC (µg/ml)	28.32^{a}	33.36^{ab}	38.96^{b}	40.81^{b}	1.84^{S}	0.009

$^{a-b}$Means bearing different superscript in the same row differ significantly($P < 0.05$); T_1 = Control, T_2 = 30 ml of *M. oleifera* leaf extract (MOLE) /kg body weight, T_3 = 60 ml of *M. oleifera* leaf extract (MOLE) /kg body weight, T_4 = 90 ml of *M. oleifera* leaf extract (MOLE) /kg body weight; SEM, Standard error of means. NS = No significant different; S = Significant different; MDA = Lipid peroxidation; TAC = Total antioxidant capacity; K = potassium level.

oleifera on haematology, serum biochemical indices and oxidation status of rabbits reared in a semi-humid environment. Results of the present study showed clearly that oral administration of MOLE to rabbits affected significantly some of the studied traits. According to Oyedemi et al. (2011), the assessment of haematological parameters could be used to reveal the deleterious effect of some chemicals in plant extracts on the blood constituents of animals. It also reflects the physiological responsiveness of the animal to its internal and external environments (Esonu et al., 2001). Observation showed that MOLE had no significant influence on the Hb, RBC and PCV values of rabbits given 30 and 90 ml MOLE/kg body weight. This shows that MOLE can be used for rabbits under heat stress without any pathological deviation from the normal. However, the recorded values for rabbits in T3 showed as a depression which perhaps could be as a result of other factors different from the experimental materials.

The present findings indicated that environmentally-induced heat stress increased the level of lipid peroxidation as reflected by the high value of malondialdehyde (MDA) in rabbits under treatment 1(0 ml of MOLE/kg body weight), in addition, heat stress also depleted the antioxidant capacity of the birds as recorded in the low level of TAC (Ismail et al., 2013). *M. oleifera leaf extract* (MOLE) decreased the incidence of lipid peroxidation as shown by the decreasing MDA values in treatments 2 (30 ml MOLE/kg body weight), 3 (60 ml MOLE/kg body weight) and 4 (90 ml MOLE/kg body weight). This is in agreement with the result of Luqman et al. (2012) who observed a lowered MDA level in mice given aqueous fruit extract of *M. oleifera*. Increased activity of antioxidant enzymes was observed in rabbit group given 30, 60 and 90 ml of MOLE/kg body weight in a dose-dependent manner. This result is similar with the results obtained by Luqman et al. (2012) who recorded total antioxidant capacity of *M. oleifera* extracts which increased with an increase in MOLE concentration. It can therefore be concluded that administration of MOLE at 90 ml /kg body weight can be used to enhance the antioxidant capacity of rabbits reared in a semi-humid

environment without any adverse effect on the blood and serum biochemistry of the test animals. Further studies can be done to evaluate use of MOLE at a higher concentration.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests**.**

REFERENCES

Baydar NG, Özkan G,Yasar S (2007). Evaluation of the antiradical and antioxidant potential of grape extracts. Food Control 18:1131-1136.

Buege JA, Aust SD (1978). Microsomal Lipid Peroxidation. Methods Enzymol. 52:302-310.

Dada AA, Abiodun AD (2014). Effect of dietary fluted pumpkin (*Telfairia occidentalis*) extract on growth performance, body composition and haematological parameters of Nile tilapia (*Oreochromisniloticus* Linnaeus).J. Fish. 2(3):203-208.

Esonu BO, Emenalom OO, Udedibie ABI, Herbert U,Ekpor CF, Okolie IC, Iheukwumere FC (2001). Performance and Blood Chemistry of Weaner Pigs Fed Raw *Mucuna* (Velvet Bean). Trop. Anim. Prod. Investig. 4:49-54.

Ewuola EO, Egbunike GN (2008). Haematological and serum biochemical growing rabbit bucks fed dietaryfumonisin. Afr. J. Biotechnol. 7(23):4304-4309.

Ganesan S, Bhatt RY (2008). Qualitative Nature of Some Traditional Crude Drugs Available in Commercial Markets of Mumbai, Maharashtra, India. Ethnobot. Leafl. 12:348-360.

Ismail IB, Al-Busadah KA, El-Bahr SM (2013). Oxidative Stress Biomarkers and Biochemical Profile in Broilers Chicken Fed Zinc Bacitracin and Ascorbic Acid under Hot Climate. Am. J. Biochem. Mol. Biol. 3:202-214.

Jaouad ElH, Zafar HI, Badiˆaa L (2004). Acute and chronic toxicological studies of *Ajugaiva*in experimental animals. J. Ethnopharmacol. 91:43-50.

Luqman S, Srivastava S, Kumar R, Maurya AK, Chanda D (2012). Experimental Assessment of Moringaoleifera Leaf and Fruit for Its Antistress, Antioxidant, and Scavenging Potential Using *In Vitro* and *In Vivo* Assays. Evid. Based Complement. Alternat. Med. Article ID 519084, 12 p.

Mitruka HM, Rawnsley SK (1997). Chemical, Biochemical and Haematological Reference Values in Normal Experimental Animals. Masson Publishing, New York, USA. pp. 287-380.

Mwale M, Masika PJ, Materechera SA (2014). Effect of medicinal planrs on haematology and serum biochemical parameters of Village chickens naturally infected with *Heterakis gallinarum.* Bangl. J. Vet. Med. 12(2):99-106.

Oduola T, Popoola GB, Avwioro OG, Oduola TA, Ademosun AA, Lawal MO (2007). Use of *Jatrophagossypifolia*stem latex as a haemostatic agent: how safe is it? J. Med. Plants Res. 1(1):14-17.

Ogbe AO, John PA (2011/12). Proximate study, mineral and anti-nutrient composition of Moringa oleifers leaves harvested from Lafia, Nigeria: Potential benefits in poultry nutrition and health. Journal of Microbiology, Biotechnology and Food Sciences 1(3) 296-308.

Ojo OA, Tanimowo DA, Oparemi AA (2015). Effects of *Moringa oleifera* leave extract on the reproductive characteristics of heat-stressed cockerels. Proceeding of 5th Nigeria, international Poultry Summit. 5th NIPS, pp. 237-244.

Oleforuh-Okoleh VU, Olorunleke SO, Nte IJ (2015). Comparative Response of Bitter Leaf (*Vernonia amygdalina*) Infusion Administration on Performance, Haematology and Serum Biochemistry of Broiler Chicks. Asian J. Anim. Sci. 9:217-224.

Oyedemi SO, Adewusi EA, Aiyegoro OA, Akinpelu DA (2011). Antidiabetic and haematological effect ofaqueous extract of stem bark of *Afzeliaafricana (Smith)* on streptozotocin-induced diabetic Wistar rats. Asian Pac. J. Trop. Biomed. 1(5):353-358.

SAS Institute (2003). SAS/STAT. User's Guide. Version 6, 4th Edition. Vol. 1 and 2. SAS Institute Inc. Cary NC.27513 USAS.

Simaraks S, Chinrasri O, Aengwanich W (2004). Haematological, electrolyte and serum biochemical values ofthe Thai indigenous chickens (*Gallus domesticus*) in north eastern, Thailand. Songklanakarin J. Sci. Technol. 26(3):425-430.

Sodamade A, Bolaji OS, Adeboye OO (2013). Proximate Analysis, Mineral Contents and Functional Properties of Moringa oleifera Leaf Protein Concentrate. IOSR Journal of Applied Chemistry, 4: 7-51.

Swanepoel C, Blockman M, Talmud J (2008). Nephrotoxins in Africa. In. de Broe ME, George AP, Bennett WM and Deray G (Editors). Clinical nephrotoxins: Renal injury from drugs and chemicals. Part 3. Metapress, Springer, New York. pp. 859-870.

academicJournals

Vol. 16(25), pp. 1391-1399, 21 June, 2017
DOI: 10.5897/AJB2016.15777
Article Number: 0FAA49864868
ISSN 1684-5315

http://www.academicjournals.org/AJB

African Journal of Biotechnology

Full Length Research Paper

Heavy metal contamination in soils from a municipal landfill, surrounded by banana plantation in the eastern flank of Mount Cameroon

Fonge, B. A.[1]*, Nkoleka, E. N.[2], Asong, F. Z.[2], Ajonina, S. A.[2] and Che, V. B.[3]

[1]Department of Botany and Plant Physiology, Faculty of Science, University of Buea, P. O. Box 63 Buea, Cameroon.
[2]Department of Environmental Science, Faculty of Science, University of Buea, P. O. Box 63 Buea, Cameroon.
[3]Department of Geology, Faculty of Science, University of Buea, P. O. Box 63 Buea, Cameroon.

Received 12 November, 2016; Accepted 4 May, 2017

Municipal solid waste generated in our cities, with an unprecedented population growth, has resulted in degrading environmental quality, thus a major problem for policy makers. The extent of Pb, Cu, Zn, Cd, Hg and Zn contamination in soils of the studied area, where vegetables were grown, using metal contamination factor (CF) and pollution load index were assessed. The concentrations of heavy metals studied were in the order Cu > Zn > Cd > Hg > Pb > Ni, with the highest value (in mg/kg) for Zn (14.15±0.73), Cu (14.15±1.59), Cd (6.57±1.71) and Hg (6.29±0.97) recorded in site SS1. The geo-accumulation index (*Igeo*) indicated that sites SS3, SS4 and SS5 were uncontaminated, moderately contaminated (1< *Igeo*<3) for Zn, landfill was moderately to heavily contaminated (1< *Igeo*<3) for Cu, Cd and for swamp only Hg. Contamination factor (CF) for soils indicated that site SS1 had a very high degree of contamination (CF > 6) for Cu and Cd while the swamp, old dumpsite and downstream had a low contamination (CF<1). The landfill area is moderately contaminated and considering the age and other anthropogenic factors, this environment may become highly polluted in future. This present work could serve as a landmark for contemporary research in eco-toxicology.

Key words: Heavy metals, contamination, soils, agriculture, landfill, Buea.

INTRODUCTION

Once released into the environment, heavy metals are often considered as a problematic environmental pollutant because of their well-known effects on living organisms. The term “heavy metal” is generally used to describe a group of metals and metalloids with an atomic density greater than 5.0 g/cm^3 and is toxic or poisonous even at low concentration (Duffus, 2002; Lenntech, 2004). Their effects on living organisms generally results from contamination of either abiotic systems (soil, water and air) and subsequent uptake and bio-accumulation/

*Corresponding author. E-mail: ambofonge@yahoo.com.

magnification by living organisms. The effects of heavy metals on the health of living organisms can be felt even when in small amounts. Heavy metals are among the leading health concerns all over the world because of their resulting long-term cumulative health effects (Huton and Symon, 1986). Heavy metals have been reported to produce mutagenic, teratogenic, neurotoxic and carcinogenic effects even at very low concentrations (Das et al., 1990; Al Saleh et al., 1996; Waalkes et al., 1999; Ngole and Ekosse, (2012; Ngole, 2015). The agency for toxic substances and disease registry (ATSDR, 2015), reported that cadmium and lead are carcinogenic, and that prolonged exposure to low concentrations of cadmium could lead to kidney disease, lung damage, fragile bones, and of Lead, nervous disorder. Heavy metals possess certain chemical and physical properties for which their use in certain products is almost unavoidable though some of such uses have been on a decline, as in the case for mercury (Aucott, 2006). When heavy metals are incorporated into products and these products are subsequently disposed of in landfills at their end-of-life or -use, there is a high possibility that with time, they will be released into the surrounding ecosystems mainly soil and water.

Increased concern on soil contamination by heavy metals has been shown in recent years. Soils are sources of substrate nutrients and are the basis of sustenance to livelihood (Abdullah et al., 2009; Asongwe et al., 2014). Thus, soils play an important role in ecological stability. Nevertheless, their quality with regards to the concentrations of heavy metals may be compromised. The origin of heavy metals in the environment could be both from natural or anthropogenic sources which include: atmospheric deposition, vehicular emission, sewage, irrigation, mining activities, industrial activities, waste disposal and agricultural applications (Asaah et al., 2006; Zhang et al., 2011; Ngole and Ekosse, 2012). They are notorious when they bio-accumulate in soil and due to their long persistence time in the course of interaction with soil component, they consequently enter food chain through plants or animals (Dosumu et al., 2003). Thus, plants grown on contaminated soils bio-accumulate these heavy metal contaminants which pose high risk to human health.

Waste disposal and agriculture are two of the anthropogenic activities that have contributed to increased levels of heavy metals in soils (Modaihsh et al., 2004; Ngole and Ekosse, 2012; Bitondo et al., 2013). Until the 1980s, the problem of municipal solid waste management (MSWM) for many communities was viewed from the cost perspective hence little attention was paid to it (Bhide and Sundersand, 1983). More so, in developing countries, MSWM was partly paralyzed by the problem of insufficient public and private funds to sustain existing schemes and corrupt management systems (Gupta et al., 1998; Buenstro et al., 2001a) and also mainly due to the fact that municipal solid waste management is influenced by a complex interrelationship of political, legal, socio-cultural, environmental and economic factors as well as available resources (Kumar et al., 2005). MSW, however, constitutes a serious environmental problem with varying degrees of direct as well as indirect negative effects on the environment and its ecosystems. Its handling across the different functional elements requires greater attention as it raises concerns not only about cost but also about environmental health and pollution. Within the EU for example, the presence of heavy metals in waste as a result of their uses in modern society is a matter of ever-growing concern to both politicians, authorities and the public (European Commission, 2002). Such concerns are slowing receiving widespread attention in developing countries like Cameroon where the fate of most MSW is disposal on open dumps and landfills. This widespread attention constitutes fall outs from the Cameroonian government's commitment to handling environmental related issues with the councils playing a very significant role in waste management. In developing countries, municipal solid waste management has most often been the responsibility of the government and/or municipalities. In Cameroon, it is the sole responsibility of municipal councils to manage MSW within the council areas, which may be through a contracted agent. Change in lifestyle and consumption habits are of particular significance to the type and quantity of municipal solid waste generated. With the usage of electrical and electronic devices on the rise (UNEP, 2009), the amount of electrical and electronic waste (e-waste) incorporated each day among the municipal solid waste (MSW) is equally growing enormously around the world. A majority of this alongside the MSW ends up in landfills and open dumps with resulting environmental and human consequences. Most noxious components including toxic substances such as heavy metals found in MSW often times leached into soil and subsequently pollute ground water at varying degrees.

Of recent, the government of Cameroon has equally given considerable attention to food insecurity with a need to boost agricultural productivity, so as to meet up with food demand of an ever increasing population. The application of pesticides and fertilizers are important inputs for agricultural production (Zhang et al., 2011), most of which may contain some amount of heavy metal in the formulae. Surrounding the Mussaka landfill is a banana plantation where there is the regular application of some pesticides and fertilizers in order to control pests/diseases as well as enhance production. The landfill receives unsorted and untreated municipal waste mainly from the following sources: Residential, commercial, institutional, medical and construction sources. The hydrogeological setting of the landfill is not the best. More so, the landfill lacks all side-sealing

systems. Thus, the effects of agricultural activities and landfill operations and reactions represent a potential source of heavy metal contamination around the Mussaka area.

There is still a dearth of studies on heavy metal pollution in Cameroon, particularly those related to soil pollution by landfills and the subsequent management of such pollution. There are a number of previous research works on municipal solid waste management in Cameroon including aspects of: generation, characterization (Achankeng, 2003; Manga et al., 2011); options of recycling and recovery (Asong, 2010). Also, some preliminary studies have been carried out on soil pollution from industries in Cameroon, mainly around Douala by Asaah et al. (2006). However, there is the need to fully investigate the impacts of waste in landfills and agricultural inputs on the surrounding soil. In Cameroon, there is no guideline for management and control of soil pollution drawn for the country but guidelines drawn from international treaties are applied. It is for this reason that baseline data has to be collected to enhance policy and to build up soil pollution guidelines.

The Mussaka landfill, that has just been operational for about five years, and located in a very moist and hot climatic setting, continues to receive huge amounts of unsorted and non-pretreated waste from Buea and its environs. The surrounding land use practices (like plantation agriculture, vegetable production and car wash unit) and the presence of sensitive ecosystems (a flowing stream, used for drinking and irrigation and a wetland) further renders this area very important as a production sites for agriculture. However these changes in land use practice have degraded the surrounding ecosystem with likely negative impacts on environment. The risk of contaminants accumulating in soil, environment and crops due to leachate from the landfill, fertilizer and pollutants is of serious concern. Given the above situation, it is very essential that the concentration of heavy metals and their potential health risk is assessed in-order to formulate policies and prevent health risk in this agricultural bread basket region. The main objectives of the current study were: (1) to determine the concentration of Pb, Cu, Zn, Cd, Hg and Zn in soils around the Mussaka landfill surrounded by plantation agriculture, and (2) delineate the extents of heavy metal contamination and their potential health risk.

MATERIALS AND METHODS

Study area description

Mussaka is located between latitude 04°08.036′ and 04°12.627′ North and Longitude 009°13.104′ and 009°18.675′ East, in the outskirt of Buea. Buea is the headquarters of South West Region and covers an area of 870 km^2, with a population of approximately 200,000 inhabitants (BUCREP, 2005), it is located on the eastern slopes of Mount Cameroon (Figure 1). The Mussaka landfill spans an area of about 13240 m^2 (1.3 ha) and is surrounded mainly by the Cameroon Development Cooperation (CDC) banana plantation. The landfill is the only current and official landfill that serves the entire Buea municipality. It receives about 104 tonnes of waste daily and has been used by HYSACAM Company for five years.

The area has a gentle to undulating relief, with a swamp and a nearby stream. The geologic setting of the landfill is volcanic, with the rock type being mainly a basaltic tuff. In vertical sections of approximately 1.5 m on some side cutting within the landfill, one finds some form of layering. The layers are marked by basalts of pebbles to cobble sizes in a mud matric being intercalated with a thin dominantly mud layer. It is because of the nature of the matrix materials and the mud that the landfill operators make use of the basaltic tuff as top sealing material. In other to meet up quantity of top sealing material, weathered tephra tuff is transported from a nearby locality and used on the landfill.

The nearby stream, flowing through the northern and eastern parts of the landfill is being used upstream for car washing and downstream for drinking and irrigation of banana, tomato and vegetable farms by the CDC Company and the locals around the area. The stream is separated from the active waste disposal front by a wetland thus rendering the area ecologically sensitive. The currently used method of disposal is a fairly controlled one characterized by: Weighing of truck load upon arrival, partitioning of landfill area into active and passive disposal spaces, spread of waste over broader areas, volume reduction by compaction with heavy duty machinery, spreading of sealing materials. These actions, although playing a role in the daily management of the landfill, fall short of appropriate measures to be taken to prevent waste in the landfill from interacting with the environment. For example, cover sealing material is placed in an extensive horizontal pattern which may create opportunity to wider migration of leachates and interaction of materials unlike in the case of creating cells of smaller sizes and concave surfaces. In terms of quantity, about a 104 tonnes of unsorted waste come into the landfill daily.

Sampling of soils

Soil samples for analyses were collected from the landfill and its surrounding area (Figure 2). Using an auger, soil samples were collected at different locations in the landfill at depths between 0 and 30 cm. The sampling sites were chosen based on the different anthropogenic activities along the landfill zone. A total of five sampling sites were identified in this study. From each site, three soil samples (in replicates) weighing an average of 900 g were collected and making a total of 15 samples analysed. Each sample was air-dried, sieved through a 2 mm sieve, and used for subsequent. The soil pH was determined using a 1:2.5 soil solution ratio with the help of the PD 300 series pH meter. The electrical conductivity was measured using the 1:5 soil solutions, where 10 g of soil was put into a beaker followed by the addition of 50 ml distilled water. The mixture was agitated for 1 h and filtered after 16 h. The value of electrical conductivity was read from a conductivity meter calibrated with 0.01 N KCl. The soil texture was determined using the hydrometer method. Soil particle size was determined using the Pipette method. The weight percent (wt %) of sand, silt, and clay was plotted in a textural triangle to determine the texture of the samples.

The following parameters were determined based on Pauwels et al. (1992): Organic carbon (OC) by the Walkley and Black method using acidified potassium dichromate; soil bases and cation exchange capacity (CEC); total nitrogen content by Kjeldahl method (TKN); concentration of exchanged aluminium is determined by calorimetry; available phosphorus was determined by the Bray II method. Heavy metals concentrations in the soil samples were

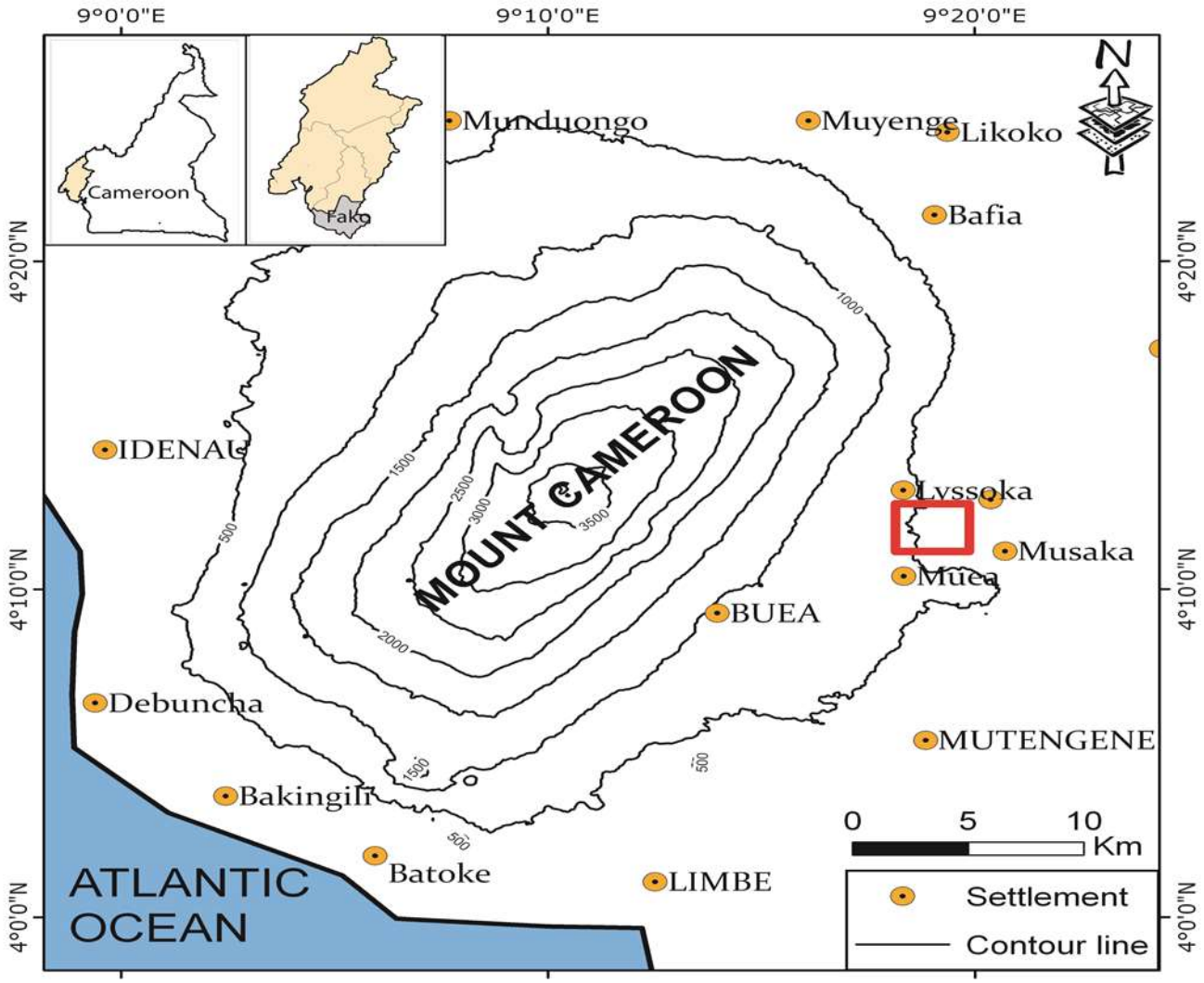

Figure 1. Location of the Mussaka landfill.

determined using atomic absorption spectrometer (Perkin Elmer). The concentrations of Zn, Cu, Ni, Pb, Cd and Hg were analyzed after acid digestion. All samples were analyzed in duplicate and the mean value reported as the concentration value for the metal.

Data interpretation

The extent of Zn, Cu, Ni, Pb, Cd and Hg contamination in soils within the Mussaka landfill vicinity was assessed using: The geo-accumulation index (I_{geo}) and the contamination factor (CF) proposed by Muller (1969) and Hakanson (1980). Some recent studies (Ngole and Ekosse, 2012) have used I_{geo} and CF to evaluate heavy metal contamination in terrestrial environment. The I_{geo} and CF were determined using the mathematical formulae as indicated in Equations 1 and 2 respectively.

Equation 1: Geo-accumulation index:

$$I_{geo} = log_2(C_n / 1.5xB_n)$$

Where, C_n = average concentration of heavy metal measured in the soil; B_n = average geochemical background concentration of the same heavy metal; 1.5 = background matrix correction factor due to lithogenic and anthropogenic influences.

Equation 2: Contamination factor:

$$CF = C_m/B_m$$

Where, C_m = measured concentration of heavy metal in the soil and B_m = local background concentration value of the heavy metal.

The concentrations of the heavy metals in the control samples (Site SS_5, control) were used as the background concentration values to calculate the heavy metal contamination factor (I_{geo} and CF) in this study. Interpretation of I_{geo} and CF were explained according to the classes as described by Ngole and Ekosse (2012).

The pollution load index (PLI) was also employed to assess the extent of heavy metal pollution in the soils. This was done with the formula used by Ngole and Ekosse (2012) as indicated in Equation 3.

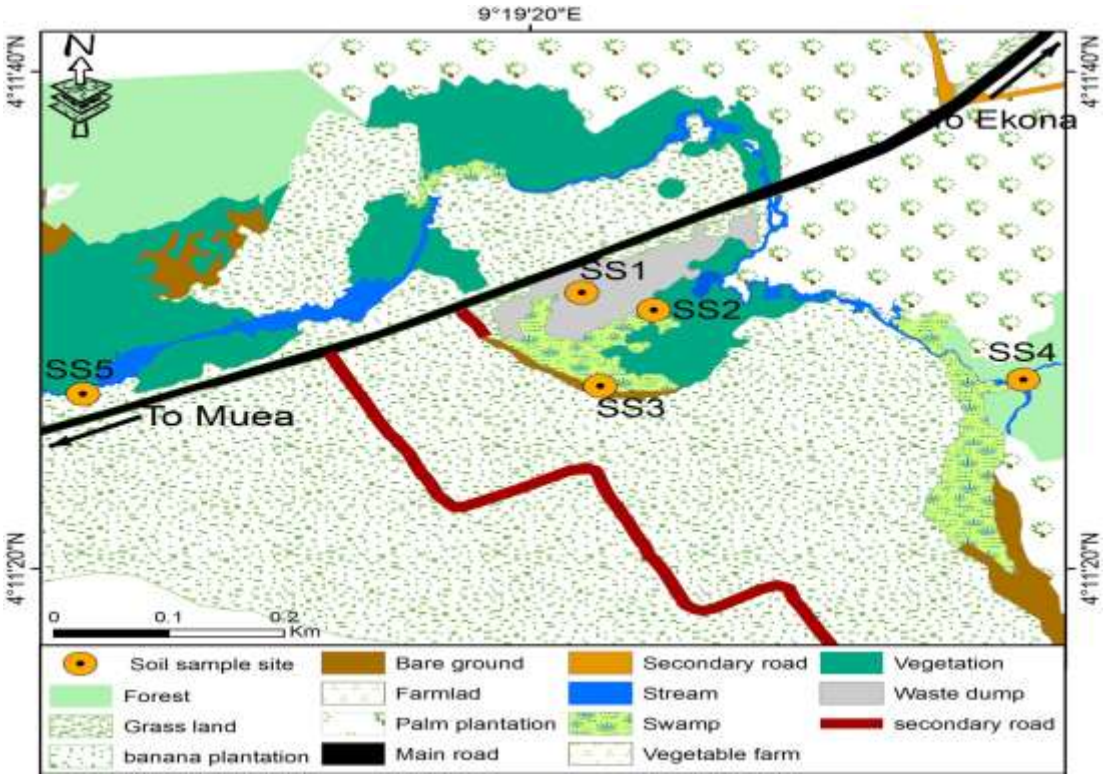

Figure 2. Study area showing sample collection points and different land use forms around the study area.

Equation 3: Pollution load index:

$$PLI = \sqrt[n]{(CF1 \times CF2 \times CF3 \times \ldots\ldots \times CFn)}$$

Where, CF is contamination factor and N is the number of elements. PLI values <1 indicate no pollution whereas values >1 indicate pollution.

RESULTS AND DISCUSSION

Soil physico-chemical properties

A number of physico-chemical properties of soils in this study were determined (Table 1). With respect to texture, the percentage of clay ranged from 1.33 to 17.67%; silt 16.67 to 29.33%; sand 53.33 to 82.00%. The percentage composition of these three components revealed that the soils in the study area could be texturally classified as sandy loam to loamy sand. This indicated a higher sand content than silt and clay. The soil physico-chemical parameters show varied correlations with heavy metals. Soil textural parameters show a slightly positive correlation with heavy metals. Soils of this nature are reported by Dube et al. (2001), to cause less dispersion of contaminants because of their high porosity and permeability.

The pH of the area ranged from moderately acidic (5.90) to slightly acidic (6.97). Heavy metals are generally more mobile at pH < 7 than at pH > 7. This can therefore be hazardous for agricultural purposes since crops are known to take up and accumulate heavy metal from contaminated soils in their edible portions (Wei et al., 2005). Such weak acidic pH in the study area can be attributed the landfill activity and agricultural inputs such as pesticides, fungicides and fertilizers. This finding is consistent with those of Srivastava (2012), who attributed acidic pH of soils to the presence of metal scrap, waste materials in the dumpsites and other anthropogenic activities at Allahabad, India.

Total nitrogen ranged from 3.5 to 4.4 gkg^{-1}. The organic carbon as well as the organic matter in soils did not vary significantly (P>0.05) among the sites, and could be due to agricultural activities in the entire study site. However, higher values of organic carbon and organic matter were recorded within the landfill sites than the control site upstream. This could be due to high organic waste input into the landfill undergoing decomposition. The organic matter plays an important role in soil structure, water retention, cation exchange and in the formation of complexes (Alloway and Ayres, 1997). The mean concentrations of Ca, Mg, K, and Na across the sites (Table 2) showed no significant differences (P > 0.05).

Table 1. Soil physico-chemical parameters for studied area.

Property	Site				
	SS_1	SS_2	SS_3	SS_4	SS_7
Clay (%)	10.36[b]	10.00[b]	17.67[ab]	1.33[a]	1.33[a]
Sand (%)	60.30[a]	65.00[a]	53.33[a]	82.00[a]	81.67[a]
Silt (%)	29.33[a]	25.00[a]	29.00[a]	16.67[a]	17.00[a]
OC (%)	3.84[a]	4.60[a]	2.33[a]	2.21[a]	1.46[a]
OM (%)	6.63[a]	7.94[a]	4.02[a]	3.81[a]	2.51[a]
Soil textural class	SL	SL	SL	LS	LS
Total nitrogen (gkg^{-1})	4.2[a]	4.4[a]	3.9[a]	3.7[a]	3.5[a]
BD (g/cm^3)	0.86[a]	ND	0.97[a]	0.97[a]	0.57[a]
C/N ratio	8.99[a]	10.45[a]	6.06[a]	5.52[a]	4.06[a]
pH_{KCl}	4.43[a]	6.40[a]	5.47[a]	5.43[a]	5.97[a]
pH_{Water}	6.97[a]	6.90[a]	6.43[a]	5.90[a]	6.57[a]
EC (mS/cm)	0.03[a]	0.04[a]	0.04[a]	0.04[a]	0.03[a]
CEC (cmol(+)/Kg)	15.17[a]	16.20[a]	10.97[a]	11.73[a]	9.61[a]
Available Bray 2-P (mg/kg)	124.36[a]	19.68[c]	28.33[b]	9.88d	26.32[b]

Means that do not share a letter within the column are statistically different.

Table 2. Mean concentrations of exchangeable cations across the sampling sites.

Site	Exchangeable cations ($cmol_+/kg$)			
	Ca^{2+}	Mg^{2+}	K^+	Na^+
SS_1	**1.47±1.11[a]**	**0.53±0.30[a]**	0.04±0.00[a]	0.03±0.00[a]
SS_2	0.24±0.00[a]	0.20±0.00[a]	0.04±0.00[a]	0.030±0.00[a]
SS_3	0.43±0.05[a]	0.24±0.05[a]	0.40±0.00[a]	0.03±0.00[a]
SS_4	0.37±0.03[a]	0.19±0.05[a]	0.40±0.00[a]	**0.02±0.00[a]**
SS_7	0.32±0.00[a]	0.03±0.03[a]	0.40±0.00[a]	0.030±0.00[a]

SS_1, Dumpsite; SS_2, Swamp below the dump; SS_3, old dumpsite; SS_4, downstream; SS_7, control site. Kruskal Wallis test was used to test for significance and the means were separated using Turkey method at $\alpha = 0.05$; Means that do not share a letter within the column are statistically different.

The mean concentrations of soil nutrients compared to critical values were as follows: Calcium was very low, magnesium was generally very low, potassium was very low and sodium was very low.

Though the concentrations of exchangeable Ca and Mg in the soils did not show any significant difference among the sites, the values within landfill were slightly higher than those of the control site. This may be connected to the heterogeneous nature of wastes received by the landfill, which is expected to impact differently on soil properties (Adjia et al., 2008).

Heavy metal concentrations in soils and risk assessment indices in study area

Heavy metals concentrations in the soils

The mean concentrations of heavy metal (in mg/kg) in soil samples collected from the study area are presented in Table 3. Within the currently active landfill zone, measured concentrations of heavy metals were in the order Zn > Cu > Cd > Hg > Ni > Pb. The concentration of zinc reveals a decreasing trend from the active landfill to the peripheries. The concentrations of the other metals do not show any trends in variations across sample sites. Compared to the EU and WHO values, most of the heavy metal values in the soil were below the permissible limits. This may be ascribed to their continuous removal by vegetables and other plants growing in the studied area. However, copper and cadmium recorded some values that were higher than WHO permissible limits. This could be attributed to leachate migration from the decomposing waste within active landfill, and accumulation from agricultural inputs from the plantation. Adjia et al. (2008) reported similar results on heavy metals content in soil of sites around landfill used for periurban agriculture in

Table 3. Mean heavy metal concentrations in soils at the Mussaka landfill vicinity compared to the European Union and WHO permissible limits.

Site	pH_{Water}	Heavy metals concentrations (mg/kg)					
		Zn	Cu	Ni	Pb	Cd	Hg
SS_1	6.97[a]	14.15±0.73[a]	14.15±1.59[a]	1.85±0.17[ab]	[BD]L	6.57±1.71[a]	6.29±0.97[a]
SS_2	6.90[a]	11.18±0.00[a]	0.76±0.00[bcd]	1.75±0.00[a]	0.49±0.00[a]	0.41±0.00[b]	3.55±0.00[a]
SS_3	6.43[a]	7.80±0.20[ab]	0.54±0.04[bcd]	2.12±0.03[b]	1.29±0.09[a]	0.73±0.06[b]	0.85±0.29[a]
SS_4	5.90[a]	8.73±0.45[ab]	1.04±0.22[abcd]	2.03±0.02[ab]	1.23±0.03[a]	0.90±0.06[ab]	1.45±0.42[a]
SS_7	6.57[a]	6.29±0.03[b]	1.68±0.23[acd]	2.15±0.01[b]	1.21±0.12[a]	0.92±0.06[ab]	1.10±0.03[a]
EU*	5<pH<6	60	20	15	70	0.5	NA
	6<pH<7	150	50	50	70	1.0	
WHO	NA	50	4	68	20	0.3	NA

SS_1, Dumpsite; SS_2, swamp below the dump; SS_3, old dumpsite; SS_4, downstream; SS_7, control site; BDL, below detectable limit; NA = not available. *Maximum permissible levels according to the EU Directive (European Union, 2002).

Table 4. Geo-accumilation index for studied heavy metals in soils around the Mussaka landfill.

Site	I_{geo}-Zn	I_{geo}-Cu	I_{geo}-Ni	I_{geo}-Pb	I_{geo}-Cd	I_{geo}-Hg
SS1	0.58	2.49	-0.81	BDL	2.25	1.93
SS2	0.24	-1.74	-0.89	-1.89	-1.74	1.10
SS3	-0.27	-2.25	-0.58	-0.49	-0.92	-0.94
SS4	-0.11	-1.29	-0.67	-0.56	-0.62	-0.18
SS7	-0.59	-0.58	-0.58	-0.60	-0.58	-0.58

SS_1, Dumpsite; SS_2, swamp below the dump; SS_3, old dumpsite; SS_4, Downstream; SS_7, control site; BDL, below detectable limit; NA = not available. I_{geo}< 0=uncontaminated, 0 < I_{geo}< 1=moderately contaminated, 1 < I_{geo}<3=moderately to heavily contaminated, 3 < I_{geo}< 4=heavily contaminated, 4 < I_{geo}< 5=heavily to extremely contaminated.

Ngaoundere. Also, similar results have been obtained by Ngole and Ekosse (2012) who reported higher concentration levels of heavy metals at sites where leachate plumes were found within the Gaborone landfill environment. Modaihsh et al. (2004), equally reported high cadmium concentrations in soils due to application of inorganic fertilizers. Cadmium is one of the most eco-toxic metals, with highly undesirable effects on soil health, plant metabolism, humans and animal health (Kabata-Pendias, 2000). At very low concentrations, chronic exposure to cadmium can result to anemia, anosmia, cardiovascular diseases and renal problems (Sharma et al., 2006).

Geo-accumulation indices heavy metals in the soils

The geo-accumulation index (*Igeo*) for the quantification of heavy metal accumulation in the soils showed that safe for the active landfill site that is moderately to heavily contaminated for most of the analysed heavy metals, all other sites were uncontaminated (Table 4). The active landfill was moderately to heavily contaminated for copper, cadmium, mercury, and zinc but uncontaminated for nickel.

This is likely influenced by leachate from the landfill, which suggests a very stern anthropogenic influence which is not unconnected with the landfill receiving all kinds of mixed wastes ranging from domestic, electronic, commercial and medical wastes (Aboyade, 2004). Zinc for instance, is a component of paint pigments, steel products, metal, automotive parts, roofings, and packaging materials (Alloway, 1995) which were components of the municipal waste deposited on the site.

Contamination factor of heavy metals in the soils

Values of the contamination factor (CF) for soils in the study area indicated that active landfill site had a high to very high degree of contamination (CF > 5) with regards to mercury, cadmium and copper. The contamination factor for heavy metals for all other sampling sites was generally less than 3 implying low to uncontaminated sites. It was also observed that CF values for copper in sampling sites other than active landfill are very low. The

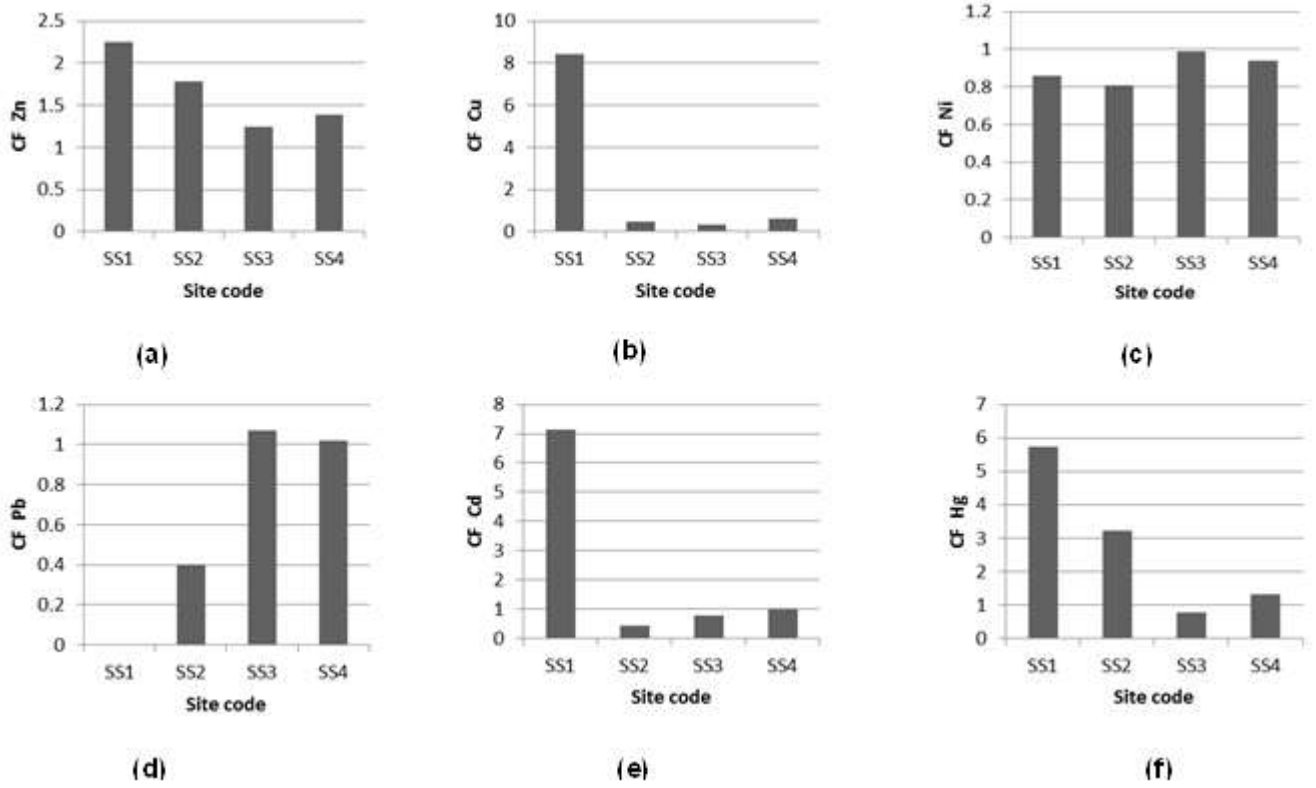

Figure 3. Contamination factor of (a) Zinc, (b) Copper, (c) Nickel, (d) Lead, (e) Cadmium and (f) Mercury in the soils within the landfill environment. (CF < 1 = low contamination; 1 < CF < 3 = moderate contaminated; 3 < CF < 6 = considerable contaminated; CF > 6 = very high contaminated).

presence of these metals could be explained as resulting from Leachate derived from landfill and agricultural origin, which are likely to contain high concentrations of heavy metals like Cu, Cd, Hg, Ni and Zn. These metals are used in manufacturing several commodities and products commonly used in homes as well as fertilizers, pesticides, herbicides and fungicides.

Pollution load index in the soils

Pollution load index (PLI) indicates the extent and infiltration of heavy metal in soil sample (Ahiamadjie et al., 2011). The pollution load index (PLI) for soils in this study took into account the combined polluting contributions of Zn, Cu, Pb, Ni, Hg and Cd (Figure 3). The values obtained for pollution index of the sampled sites within the study area indicated that Site SS_2 and SS_3 were not polluted (PLI < 1) while site SS_1 and SS_4 were polluted (PLI > 1). This can be explained by the high concentrations of Cu, Zn, Cd and Hg, which could have resulted from the decomposition of assorted waste in the landfill (site SS_1) and intense fertilizer and pesticide inputs at site SS_4.

Conclusion

In this study, an evaluation of the concentration of some heavy metal (Zn, Cu, Cd, Pb and Hg) as well as risk assessment indices was done for soils within the Musaka municipal solid waste landfill surrounded by a banana plantation. Although, most of the heavy metal concentrations in the soil were below internationally permissible limits, elevated levels of Cu and Cd in the soils at certain sites suggests possible need for concern with respect to future pollution as the landfill maturation progresses. The geo-accumulation and pollution indices of soil further reveal that with the current state of activity, the landfill is contaminated with heavy metals. Given that this landfill is relatively young, there is a need to develop monitoring program so as to enable subsequent decisions on water, land, and habitat use at the periphery of the landfill as well as downstream of water bodies to be made.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

ACKNOWLEDGEMENT

The authors acknowledge and extend their sincere gratitude to all who assisted to the realisation of this present work. Special thanks go to the HYSACAM company for granting access to the landfill, to the University of Buea (Life Science Laboratory), and University of Dschang (Laboratory of Soils and Environmental Chemistry) for assisting in chemical analysis.

REFERENCES

Abdulahi MM (2009). Municipal solid waste incineration bottom ash as road construction material. Assumption Uni. J. Technol. 13(2):121-128.

Aboyade, A. (2004). The potential for climate change mitigation in the Nigerian solid waste disposal sector: a case study from Lagos. Masters thesis, Lund University, Sweden.

Achankeng, E (2003). Globalization, urbanization and municipal solid waste management in Africa. African studies association of Australasia and the Pacific, conference proceedings – Africa on the global stage.

Adjia R, Fezeu WML, Tchatchueng JB, Sorho S, Echevarria G, Ngassoum MB (2008). Long-term effect of municipal solid waste amendment on soil heavy metal content of sites used for periurban agriculture in Ngaoundere, Cameroon. African J. of Envi. Sc. and Techno. 2(12):412-421.

Ahiamadjie H, Adukpo K, Tandoh B, Gyampo O, Nyarku I, Mumuni I, Agyemang O, Ackah O, Otoo F, Dampare SB (2011). Determination of the elemental contents in soils around diamond cement factory, Aflao. Res. J. l of Env. Earth Sc. 3(1):46-50.

Alloway BJ. Soil pollution and land contamination. In: Harrison, R.M (ed.) (1995). Pollution, causes, effects and controls. The Royal Society of Chemistry, Cambridge. P 318.

Alloway BJ, Ayres DC (1997) Chemical principles of environmental pollution. Chapman and Hall, London.

Al-Saleh I, Coate L, Al-Humaidan E, Al-Sedairy S (1996). Effect of lead on the lymphocyte subsets in Saudi Arabia with special reference to smoking. Trace Elements Electrolytes. 13(2):92-96.

Asaah VA, Abimbola AF, Suh CE (2006). Heavy metal concentrations and distribution in surface soils of the Bassa industrial zone 1, Douala, Cameroon. Arabian J. Sci. Eng. 31(2a):147-158.

Asong FZ (2010) Recycling and Material Recovery in Cameroon: Implications for Poverty Alleviation and Ecological Sustainability. Catalogue of the German National Library. P 157. http://dnb. info/gnd/14202662X

Asongwe G A, Yerima BPK, Tening AS (2014). Vegetable production and the livelihood of farmers in Bamenda Municipality, Cameroon. Int. J. Current Microbiol. Appl. Sci. 3(12):682-700.

ATSDR Toxic substances portal. Toxicological profiles 2015.http:// www.atsdr.cdc.gov/toxprofiles/index.asp. (assessed February 19 2016).

Aucott M (2006). The fate of heavy metals in landfills: A review, prepared for the "Industrial Ecology, Pollution Prevention and the NY-NJ Harbor" Project of the New York academy of sciences.

Bhide AD, Sundersand BB (1983). Solid waste management in developing countries. Indian National Scientific Documentation Centre, New Delhi, India.

Bitondo D, Tabi FO, Kengmegne SSA, Ngoucheme M, Mvondo Ze AD (2013). Micronutrient concentrations and environmental concerns in an intensively cultivated typic dystrandept in Mount Bambouto, Cameroon. Open J. Soil Sci. 3:283-288.

BUCREP (2005).Résultats du rescencement Cameroun. Bureau central des recensements et des études de la population2005. http://www.bucrep.cm/en/resources-and-documentation/downloads.html

Buenstro O, Bocco G, Bernache G (2001a). Urban solid waste generation and disposal in Mexico: A case study. Waste Man. and Res. 19:169 -176.

Das D, Srinivasu M, Bandyopadhyay M (1998). Solid state acidification of vegetable waste. Indian J. Environ. Health. 40(4):333-342.

Dosumu O, Salami N, Adekola FA (2003). Comparative study of trace element levels. Bul. of the Chem. Soc. of Ethiopia. 17(1):pp107-112.

Dube A, Zbytniewski R, Kowalkowski T, Cukrowska E, Buszewski B (2001). Adsorption and migration of heavy metals in soil. Polish Journal of Environmental Studies 10(1):1-10.

Duffus JH (2002). Heavy metals – a meaningless term". Pure Appl. Chem. 74:793-807.

European Union (2002). Heavy metals in wastes, European Commission on Environment. Available from: http://ec.europa.eu/environment/waste/studies/pdf/heavymetalsreport .pdf

Gupta S, Krishna M, Prasad RK, Kansal A (1998). Solid waste management in India: options and opportunities. Resour. Conserv. Recyc. 24:137-154.

Hakanson L (1980). An ecological risk index for aquatic pollution control: a sedimentological approach. Water Res. 14(8):975-1001

Hutton M, Symon C (1986). The quantities of cadmium, lead, mercury and arsenic entering the UK environment from human activities. Sci. Total Environ. 57:129-150.

Kabata-Pendias A (2000). Trace element in soils and plants. CRC Press: Boca Raton, FL, USA. P 413.

Kumar V, Bee DJ, Shirodkar PS, Tumkor S, Bettig BP, Sutherland JW (2005). Towards sustainable product and material flow cycles: identifying barriers to achieving product multi-use and zero waste. In: Proc. IMECE 2005. 2005 ASME International Mechanical Engineering Congress and Exposition, November 5–11, Orlando, FL, USA.

Lenntech water treatment and air pollution. Water treatment, Lenntech, Rotterdamseweg. Netherlands. 2004. Available from www.excelwater.com/thp/filters/water–purification (Retrived May 14th, 2015)

Manga VE, Forton OT, Mofor LA, Woodard R (2011). Health care waste management in Cameroon: A case study from the Southwestern Region. Resour. Conserv. Recycl. 57:108-116.

Modaihsh AS, Al-Swailem MS, Mahjoub MO (2004). Heavy metals content of commercial inorganic fertilizers used in the Kingdom of Saudi Arabia. Agric. Marine Sci. 9(1):21-25.

Muller G (1969). Index of geoaccumulation in sediments of the Rhine River. J. Geol. 2:109-118

Ngole VM (2015). Heavy metals in soils along unpaved roads in south west Cameroon: contamination levels and health risks. A J. Human Environ. 45(3):374-386.

Ngole VM, Ekosse GE (2012). Copper, nickel and zinc contamination in soils within the precincts of mining and landfilling environments. Intl. J. Environ. Sci. Technol. 9:485-494.

Sharma RK, Agrawal M, Marshall F (2006). Heavy metal contamination in vegetables grown in wastewater irrigated areas of Varanasi, India. Bull. Environ. Contamin. Toxicol. 77:312-318.

Srivastava KP, Singh VK (2012). Impact of air-pollution on pH of soil of Saran, Bihar, India. Res. J. Recent Sci. 1(4):9-13.

United Nations Environmental Program (UNEP), Recycling—From E-Waste to Resources, New York: UNEP, 2009. www.unep.org (accessed on 23.01.16)

Waalkes MP, Anver MR, Diwan BA (1999). Chronic toxic and carcinogenic effects of oral cadmium in the Noble (NBL/Cr) rat: induction of neoplastic and proliferative lesions of the adrenal, kidney, prostate, and testes. J. Toxicol. Environ. Health Part A. 58(4):199-214.

Wei S, Zhou Q, Xin W (2005). Identification of weed plants excluding the uptake of heavy metals. Environ. Intl. 31:829-834.

Zhang W, Jiang F, Ou J (2011). Global pesticide consumption and pollution: with China as a focus. Proceed. Intl. Acad. Ecol. Environ. Sci. 1(2):125-144.

academicJournals

Vol. 16(25), pp. 1400-1403, 21 June, 2017
DOI: 10.5897/AJB2017.16073
Article Number: BF079EB64870
ISSN 1684-5315

http://www.academicjournals.org/AJB

African Journal of Biotechnology

Full Length Research Paper

Improvement of color and increase in anthocyanin content of 'Niagara Rosada' grapes with application of abscisic acid

Marco Antonio Tecchio[1], Francisco José Domingues Neto[1]*, Adilson Pimentel Junior[1], Marlon Jocimar Rodrigues da Silva[1], Sérgio Ruffo Roberto[2] and Ronny Clayton Smarsi[3]

São Paulo State University (UNESP), School of Agriculture, Botucatu, SP, Brazil.

Received 15 May, 2017; Accepted 15 June, 2017

This study aimed to assess the application effect of abscisic acid (*S*-ABA) on color, and total anthocyanin and phenolic compound contents of the 'Niagara Rosada' grapes. The experiment was carried out in a commercial vineyard located in Jales, SP, Brazil, during 2014/2015. The randomized block design was used with four treatments and five replications. The treatments consisted of applying different concentrations of *S*-ABA: 0 mg L^{-1} (control); 400 mg L^{-1} of *S*-ABA at the beginning of maturation (BM); 400 mg L^{-1} of *S*-ABA in the BM + 200 mg L^{-1} 25 days after the first application (25 DAFA) and 400 mg L^{-1} of *S*-ABA in IM + 400 mg L^{-1} at 25 DAFA. At harvest, the following variables were evaluated: luminosity (*L), saturation (*C**), hue (*h*°) and color index (CIRG). The total anthocyanin content, as well the total phenolic compounds of berries were also determined. The application of 400 mg L^{-1} of *S*-ABA in the BM + 200 mg L^{-1} to 25 DAFA significantly increased the total anthocyanin content of the 'Niagara Rosada' grapes, improving the color of the berries as well.**

Key words: Abscisic acid, Niagara Rosada' grapes, color, total anthocyanin, phenolic compound.

INTRODUCTION

The State of São Paulo is the second largest producer of table grapes in Brazil (Agrianual, 2015), with emphasis on the cultivation of common or hybrids grapes, such as 'Niagara Rosada'. This cultivar is a colored mutation of 'Niagara' (*Vitis labrusca* L. x *Vitis vinifera* L.), and it has a foxy flavor and scent, and still presents satisfactory sugars level that attracts consumers. 'Niagara Rosada' is cultivated mainly in the eastern region of São Paulo, and there are, however, other winegrowing regions, such as the northwest and southwest of the State of São Paulo, represented respectively by the cities of Jales and São Miguel Arcanjo.

Traditionally, the Northwest region of the São Paulo State stands out in the cultivation of table grapes (*V. vinifera*), however, a decrease in the production of these grapes has been observed in the last years and the increase in the production of common grapes (*V. labrusca* L. x *V. vinifera* L.), predominantly the 'Niagara

*Corresponding author. E-mail: fjdominguesneto@hotmail.com.

Rosada', which, during the last two years had an increase in production of 17.4% (IEA, 2017).

The climate conditions of the northwest region allow the harvesting of the 'Niagara Rosada' grapes from July to October, the offspring of the eastern and southwestern regions of the state, traditional growing regions where the harvesting of the summer crops is carried out from December to February, and the winter crops from March to June. Associated with the best price obtained in the off-season, there is also a significant reduction of labor and agricultural pesticides in the growing of 'Niagara Rosada' when compared with fine table grapes. However, the warm climate and the low thermal amplitude during the maturation stage of the grapes in the Northwest region make them not to reach optimal color levels (Gardin et al., 2012) due to the reduction of anthocyanin synthesis (Sozim, 2011). Thus, it is necessary to apply plant regulators that aid in these compounds synthesis.

Historically, the acquisition value of ABA used to be high, however, as time went by, a cheaper production method was developed, designated as *S*-ABA, which is more accessible to grape growers (Cantín et al., 2007). Several studies suggest that exogenous applications of ABA increase the anthocyanin content in the skin of colored grapes, surpassing even the ethephon, playing a vital role in the activation of flavonoids (Cantín et al., 2007; Peppi et al., 2007; Sandhu et al., 2011; Roberto et al., 2013; Koyama et al., 2014a; Domingues Neto et al., 2017).

Color is one of the decisive attributes when marketing table grapes, since those with better color uniformity receive better values because color plays a direct role in attracting consumers, therefore, it is one of the most important ampelographic features, as well as being one of the main characteristics to be considered to evaluate the quality by phenotyping of a potential genotype (Rustioni et al., 2013).

Anthocyanins, as an example of some phytochemicals of the polyphenol class, are responsible for the color on red grape berries and their accumulation seem to be regulated, at least in part, by abscisic acid - ABA (Lacampagne et al., 2010) and these pigments are considered important for the commercial value of the grapes (Rustioni et al., 2015), as well as act as antioxidants in the human body, preventing several degenerative diseases (Nixdorf and Hermósín-Gutiérrez, 2010). The expression of anthocyanins depends on internal factors, such as abscisic acid, which induces the transmission factor, MYB1A, a protein responsible for regulating the transcription of genes that make up the biosynthetic route of the anthocyanins of the colored grapes (Jeong et al., 2004).

Considering the increase of 'Niagara Rosada' grape in São Paulo State and the lack of studies involving the application of plant regulators to improve grape quality, this study had as objective to evaluate the effect of *S*-ABA application in order to increase color, total anthocyanins content and the total phenolic compounds of 'Niagara Rosada' grape.

MATERIALS AND METHODS

The experiment was carried out in a commercial 6-year old vineyard of 'Niagara Rosada' (*V. labrusca* x *V. vinifera*) grapevine, grafted on the IAC 572 'Jales' (*Vitis caribaea* x 101-14 Mgt) rootstock, located in Jales, SP (20º 16 'S, 50º 33' W and average altitude of 480 m), during 2014/2015 season. Vines were trained in an overhead trellis system, spaced at 2.50 x 2.00 m distance and irrigated by micro sprinklers.

The climate of the region according to Köppen is *Aw* (tropical climate with dry winter season), with annual average rainfall of 1221.6 mm. The annual average temperature is of 24.4°C, with a minimum average of 14.0°C and maximum of 33.0°C (CEPAGRI, 2016).

The randomized block design was used with four treatments and five replications, and each plot was composed by one vine. The treatments consisted of the application of different concentrations of the isomer *S*-ABA (100 g L^{-1} of active ingredient, Valent BioSciences Co.): 0 (control); 400 mg L^{-1} of *S*-ABA at the beginning of maturation (BM); 400 mg L^{-1} of *S*-ABA in BM + 200 mg L^{-1} at 25 days after the first application (25 DAFA) and 400 mg L^{-1} of *S*-ABA in BM + 400 mg L^{-1} at 25 DAFA. When preparing the solution for all treatments, BreakThru®, a non-ionic spreader (0.3 mL L^{-1}) was added. The beginning of maturation of the berries was when they started softening and color change was considered, with soluble solids of 10 °Brix and titratable acidity of 1.02% of tartaric acid. During all applications, the clusters were pulverized in the morning using a backpack sprayer under 40 kgf cm^2 of pressure with JA1 hollow cone jet nozzles until the berries were completely and even covered, the syrup volume employed was of 800 L ha^{-1}.

At harvest, 10 representative clusters per plot were selected, from which 100 berries were collected (10 berries per bunch), which were collected from the lower, middle and upper parts of the cluster (3:4:3) in order to determine color parameters, and total anthocyanin and total phenolic compounds content.

For color measurements, the Minolta CR-10 colorimeter was used and the variables of the berries equatorial portion were then obtained: L^* (luminosity), C^* (saturation) and $h°$ (hue angle) (Koyama et al., 2014a). The berries color index (CIRG) was determined using the formula: $CIRG = [(180 - h^o)/(L^* + C^*)]$ (Cantín et al., 2007).

For total anthocyanin assessments, berries were cut in half and the seeds were removed, and then, they were pulverized with liquid nitrogen, according to Lee and Francis (1972). The readings were performed on a Genesys 10S Spectrophotometer, UV-VIS at 535 nm. In order to calculate the total anthocyanins concentration, the formula proposed by Teixeira et al. (2008) was used. Then, the quantification of total phenolic compounds was carried out using the Folin-Ciocalteau method, according to Bucic-Kojic et al. (2007), using 50% ethanol as an extractor. The readings were performed on a Genesys 10S Spectrophotometer, UV-VIS at 765 nm. The data obtained were submitted to analysis of variance and the means were compared by Tukey test, at 5% probability.

RESULTS AND DISCUSSION

It was noticed that two applications of *S*-ABA, one performed at the beginning of maturation and the other 25 days after the first application (400 + 200 mg L^{-1} or 400 + 400 mg L^{-1}), provided lower values of luminosity

Table 1. Luminosity (*L**), saturation (*C**), hue (*h*°) and color index (CIRG) of 'Niagara Rosada' grape, subjected to the application of different concentrations of abscisic acid (*S*-ABA), Jales, Brazil, 2014/2015.

Concentrations of *S*-ABA (mg L^{-1})	*L**	*C**	*h*°	CIRG
Control	32.98 ± 1.0[a]	10.07 ± 0.8[a]	32.97 ± 8.8[a]	3.42 ± 0.3[a]
400 (BM)	32.62 ± 1.7[a]	9.71 ± 0.6[a]	35.45 ± 8.6[a]	3.42 ± 0.4[a]
400 (BM) + 200 (25 DAFA)	29.88 ± 1.8[b]	10.18 ± 0.8[a]	34.26 ± 6.1[a]	3.63 ± 0.3[a]
400 (BM) + 400 (25 DAFA)	28.60 ± 1.4[b]	10.31 ± 0.9[a]	31.72 ± 3.6[a]	3.81 ± 0.2[a]
CV (%)	4.9	7.8	23.0	14.9

Means followed by the same letter in the column do not show any significant difference by the Tukey's test at 5% probability. BM = Beginning of maturation; 25 DAFA = 25 days after the first application (DAFA).

(*L**) of berries (29.88 and 28.60, respectively), on a scale from 0 to 100 and from black to white, respectively (Table 1). There was no significant difference for *L** between control and 400 mg L^{-1} of *S*-ABA applied at the beginning of maturation, with means of 32.98 and 32.62, respectively.

Lower *L** indicate darker color, and it means that two applications of *S*-ABA provided an increase in the 'Niagara Rosada' grape color, which may reflect directly on its commercialization; once consumers are attracted to, among other factors, the color of berries, a factor that might favor its commercial value (Roberto et al., 2012; Domingues Neto et al., 2017).

There was no significant difference between the different *S*-ABA concentrations regarding chroma (*C**), which are responsible for color saturation; hue angle (*h*°) and color index (CIRG), which ranged from 9.71 to 10.31 for *C**, from 31.72 to 35.45 for *h*° and from 3.42 to 3, 81 for the CIRG. These results corroborate with those found by Roberto et al. (2012), that obtained for 'Benitaka' table grape, with two applications of *S*-ABA 200 + 200 mg L^{-1} and 400 + 400 mg L^{-1}, both at 7 days after maturation and 15 days before harvest, a decrease of *L**, presenting no changes in the *C** and *h*°. However, the CIRG found by these authors in the control treatment was lower than all treatments that underwent the application of *S*-ABA, which was not observed in this study. The results of the present study are similar to those found in studies, which also evaluated *S*-ABA, with 'Redglobe' (Peppi et al., 2007), 'Flame Seedless' (Peppi and Fidelibus, 2008) and 'Crimson Seedless' grapes.

In experiments with 'Isabel' grape (Koyama et al., 2014b), 'Benitaka' (Roberto et al., 2012) and 'Flame Seedless' (Peppi and Fidelibus, 2008) grapes, the *S*-ABA application decreased *L** and *C** values. For 'Rubi' grape, Roberto et al. (2013) found differences in the variable *h*º, but the *S*-ABA provided lower values than those of the control. Several factors can interfere with the results of these variables, such as the grape maturation, cultivar and cultural practices, among others.

The 'Niagara Rosada' grapes which received two *S*-ABA applications, regardless of their concentrations, presented total anthocyanin contents twice as high as compared to the control treatment, with average values of 0.80 and 0.40 mg 100 g^{-1}, respectively. The synthesis of anthocyanins can occur, constantly, from the beginning of maturation to harvest, as well as during the first half of maturation (Mullins et al., 1992), which justifies the positive effect of *S*-ABA application on two grape ripening stages. Several studies have demonstrated that the exogenous application of *S*-ABA increases the anthocyanin content in grapes (Ban et al., 2003; Yamamoto et al., 2015, Domingues Neto et al., 2017). It is possible to state that, with the results obtained in this study and those already available in the literature, that the *S*-ABA triggers the secondary metabolism of 'Niagara Rosada' grapes and other grape cultivars, specifically the anthocyanin biosynthesis, which is the pigment that imparts a pink/purple color on fruits, flowers and vegetables.

Despite providing a greater accumulation of total anthocyanins, the application of *S*-ABA presented no effect on total phenolic compounds content of 'Niagara Rosada' grapes (Table 2), with average values varying from 33.02 to 39.85 mg 100 g^{-1}. Similar results were found in 'Isabel' grapes, in which applications of *S*-ABA 400 + 400 mg L^{-1} (7 days after maturation and 10 days before harvest) provided an accumulation of total anthocyanins, however, with no difference in total phenolic compounds content (Koyama et al., 2014b; Yamamoto et al., 2015).

The *S*-ABA had an effect on the enhancement of color characteristics, as well as on the synthesis and accumulation of total anthocyanins of 'Niagara Rosada' grapes, especially when it was applied twice. Thus, the use of *S*-ABA is a promising alternative for viticulture, once it makes possible, addition of more value to the fruits, since the grapes with a better color uniformity get better values in the market, and 'Niagara Rosada' grapes, have in their color, one of the decisive parameters to obtain better prices.

Since there was no significant difference in the treatments with two *S*-ABA applications, 400 mg L^{-1} of *S*-ABA at the beginning of maturation + 200 mg L^{-1} 25 days after the first application can be used for color improvement, considering the use of a lower concentration during the second application, aiming to reduce the cost of the process.

Table 2. Anthocyanins and total phenolic compounds of 'Niagara Rosada' grape, subjected to the application of different concentrations of abscisic acid (*S*-ABA), Jales, Brazil, 2014/2015.

Concentrations of *S*-ABA (mg L^{-1})	Anthocyanins (mg 100 g^{-1})	Phenolics compounds (mg 100 g^{-1})
Control	0.40 ± 0.14^{b}	39.85 ± 11.4^{a}
400 (BM)	0.56 ± 0.27^{ab}	35.44 ± 4.37^{a}
400 (BM) + 200 (25 DAFA)	0.80 ± 0.32^{a}	35.10 ± 6.29^{a}
400 (BM) + 400 (25 DAFA)	0.80 ± 0.20^{a}	33.02 ± 6.34^{a}
CV (%)	37.3	21.1

Means followed by the same letter in the column do not show any significant difference by the Tukey's test at 5 % probability. BM = Beginning of maturation; 25 DAFA = 25 days after the first application (DAFA).

Conclusions

Two applications of *S*-ABA (400 mg L^{-1} of *S*-ABA at the beginning of maturation + 200 mg L^{-1} 25 days after the first application) are efficient in improving the color of berries, as well the anthocyanins accumulation of 'Niagara Rosada' grapes.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

REFERENCES

Agrianual - Anuário da Agricultura Brasileira (2015). *Uva*: produção brasileira, São Paulo, 464 p.

Ban T, Ishimaru M, Kobayashi S, Shiozaki S, Goto-Yamamoto N, Horiuchi S (2003). Abscisic acid and 2,4-dichlorophenoxyacetic acid affect the expression of anthocyanin biosynthetic pathway genes in 'Kyoho' grape berries. J. Hortic. Sci. Biotechnol. 78(1):586-589.

Bucic-Kojic A, Planinic M, Tomas S, Bilic M, Velic D (2007). Study of solid-liquid extraction kinetics of total polyphenols from grapes seeds. J. Food Eng. 81(1):236-242.

Cantín CMA, Fidelibus BMW, Crisosto CH (2007). Application of abscisic acid (ABA) at veraison advanced red color development and maintained postharvest quality of 'Crimson Seedless' grapes. Postharvest Biol. Technol. 46(3):237-241.

CEPAGRI - Centro de Pesquisas Meteorológicas e Climáticas Aplicadas a Agricultura (2016). Clima dos municípios paulistas: São Paulo, Disponível em: <http://www.cpa.unicamp.br/outras-informacoes/clima_muni_284.html>. Visited in January 29, 2017.

Domingues Neto FJ, Tecchio MA, Pimentel Junior A, Vedoato BTF, Lima GPP, Roberto SR (2017). Effect of ABA on colour of berries and in the anthocyanin accumulation and total phenolic compounds of 'Rubi' table grape (*Vitis vinifera*). Aust. J. Crop Sci. 11:199-205.

Gardin JPP, Schumacher RL, Bettoni JC, Petri JL, Souza EL (2012). Ácido abscísico e etefom: influência sobre a maturação e qualidade das uvas 'Cabernet Sauvignon'. Rev. Bras. Frutic. 34(1):321-327.

IEA - Instituto de Economia Agrícola (2017). Estatísticas da produção paulista. <http://www.iea.sp.gov.br/out/index.php>. Visited in January 29, 2017.

Jeong ST, Goto-Yamamoto N, Kobayashi S, Esaka M (2004). Effects of plant hormones and shading on the accumulation of an-and the expression of anthocyanin biosynthetic genes iberry skins. Plant Sci. 167(2):247-252.

Koyama R, Assis AM, Yamamoto LY, Borges WF, Borges RS, Roberto SR (2014a). Exogenous abscisic acid increases the anthocyanin concentration of berry and juice from 'Isabel' grapes (*Vitis labrusca* L.). Hortscience 49(4):460-464.

Koyama R, Yamamoto LY, Borges WFS, Pascholati MB, Borges RS, Assis AM, Roberto SR (2014b). Épocas de aplicação e concentrações de ácido abscísico no incremento da cor da uva 'Isabel'. Semin. Ciên. Agrar. 35(4):1697-1706.

Lacampagne S, Gagné S, Gény L (2010). Involvement of Abscisic Acid in Controlling the Proanthocyanidin Biosynthesis Pathway in Grape Skin: New Elements Regarding the Regulation of Tannin Composition and Leucoanthocyanidin Reductase (LAR) and Anthocyanidin Reductase (ANR) Activities and Expression. J. Plant Growth Regul. 28(1):81-90.

Lee DH, Francis FJ (1972). Standardization of pigment analyses in cranberries. HortScience 7(1):83-84

Mullins MG, Bouquet A, Willians LE (1992). Biology of the Grapevine Cambridge University Press, Cambridge, 252p.

Nixdorf SL, Hermosín-Gutiérrez I (2010). Brazilian red wines made from the hybrid grape cultivar Isabel: Phenolic composition and antioxidante capacity. Anal. Chim. Acta 659(1):208-215.

Peppi MC, Fidelibus MW (2008). Effects of forchlorfenuron and abscisic acid on the quality of 'Flame Seedless' grapes. Hortscience 43(1):173-176.

Peppi MC, Fidelibus MW, Dokoozlian N (2007). Application timing and concentration of abscisic acid affect the quality of 'Redglobe' grapes. J. Hortic. Sci. Biotechnol. 82(2):304-310.

Roberto SR, Assis AM, Yamamoto LY, Miotto LCV, Sato AJ, Koyama R, Genta W (2012). Application timing and concentration of abscisic acid improve color of 'Benitaka' table grape. Sci. Hort. 142(1): 44-48.

Roberto SR, Assis AM, Yamamoto LY, Miotto LCV, Koyama R, Sato AJ, Borges RDS (2013). Ethephon use and application timing of abscisic acid for improving color of 'Rubi' table grape. Pesqui. Agrop. Bras. 48(7):797-800.

Rustioni L, Basilico R, Fiori S, Leoni A, Maghradze D, Failla O (2013). Grape colour phenotyping: development of a method based on the reflectance spectrum. Phytochem. Anal. 24(1):453-459.

Rustioni L, Rocchi L, Failla O (2015). Effect of anthocyanin absence on white berry grape (*Vitis vinifera* L.). Vitis 54(1):239-242.

Sandhu AK, Gray DJ, Lu J, Gu L (2011). Effects of exogenous abscisic acid on antioxidant capacities, anthocyanins and flavonol contents of muscadine grape (*Vitis rotundifolia*) skins. Food Chem. 126(3):982-988

Sozim M (2011). Efeito da aplicação de um análogo de brassionosteróide sobre a maturação da uva 'Niagara Rosada'. Dissertação de Mestrado, Universidade Estadual de Ponta Grossa, Brasil 50p.

Teixeira LN, Stringheta PC, Oliveira FA (2008). Comparação de métodos para quantificação de antocianinas. Rev. Ceres 55(4):297-304.

Yamamoto LY, Assis AM, Roberto SR, Bovolenta YR, Nixdorf SL, García-Romero E, Gómez-Alonso S, Hermosín-Gutiérrez I (2015). Application of abscisic acid (*S*-ABA) to cv Isabel grapes (*Vitis vinifera*×*Vitis labrusca*) for color improvement: Effects on color, phenolic composition and antioxidant capacity of their grape juice. Food Res. Int. 77:572-583.

academicJournals

Vol. 16(25), pp. 1404-1427, 21 June, 2017
DOI: 10.5897/AJB2016.15860
Article Number: E56524C64872
ISSN 1684-5315

http://www.academicjournals.org/AJB

African Journal of Biotechnology

Full Length Research Paper

DEXseq and Cuffdiff approaches weighing differential spliced genes exons modulation in estrogen receptor β (Erβ) breast cancer cells

Noel Dougba Dago[1,2]*, Nafan Diarrassouba[1], Inza Jesus Fofana[1], Jean-Luc Aboya Moroh[1], Hermann-Désiré Lallié[1], Didier Martial Yao Saraka[1], Oléfongo Dagnogo[3], Souleymane Silué[1], Armel Herve Nwabo Kamdje[4], Lamine Baba-Moussa[5] and Alessandro Weisz[2]

[1]Unité de Formation et de Recherche (UFR) Sciences Biologiques, Unité Pédagogique et de Recherche de Génétique, Université Péléforo Gon Coulibaly BP1328 Korhogo, Côte d'Ivoire.
[2]Laboratory of Molecular Medicine and Genomics, Department of Medicine and Surgery, University of Salerno, Via S. Allende, 1, Baronissi, SA 84081, Italy.
[3]Unité de Formation et de Recherche (UFR) Biosciences, Laboratoire de Pharmacodynamie Biochimie, Université Felix Houphoet-Boigny BP V 34 Abidjan 01, Côte d'Ivoire.
[4]Department of Biomedical Sciences, Faculty of Science, University of Ngaoundere P. O. Box 454 Ngaoundere, Cameroon.
[5]Laboratoire de Biologie et de Typage Moléculaire en Microbiologie, Faculté des Sciences et Techniques, Université d'Abomey-Calavi, Cotonou, Benin.

Received 28 December, 2016; Accepted 9 June, 2017

While the alternative transcription and splicing mechanisms have long been known for some genes like oncogenes, their prevalence in almost all multi-exon genes has been recently realized with the increasing application of high-throughput experimental methods, named Next Generation RNA Sequencing (NGS). Henceforth, understanding the regulation of these processes in comparisons between cell types and cancer requests, sensitive and specific bioinformatics as well as bio-statistic approaches, that is, Cufflinks/Cuffdiff, DEXseq and RESEM, detecting gene transcript/isoforms and exons abundance is necessary. Isoforms and exons expression analysis by NGS is complicated by several sources of measurement variability causing numerous statistical defies. Here, with the purpose of minimizing this statistical challenge, we integrated both Cufflinks/Cuffdiff and DEXseq bioinformatics approaches assessing whole alternative splicing (AS) events, focusing on alternative transcripts regulation and their exons modulation respectively, by processing our previous prepared Estrogen Receptor β ($Erβ^+$ and $Erβ^-$) breast cancer (BC) cells, stimulated by estradiol (E_2). Results showed that Estradiol (E_2) induced $Erβ^+$ BC ($Erβ^+E_2$), exhibited dissimilar reply as opposed to the other's analyzed BC cell lines in terms of intragenic, exons and junction reads count ratio. Relationship analysis between expressed genes and transcript isoforms, suggested a substantial role of alternative promoters in AS event occurrence in $Erβ^+$ BC as opposed to $Erβ^-$ BC. Indeed, merging Cufflinks/Cuffdiff and DEXseq approaches, 79 multi-exon genes were detected as statistically differentially modulated (spliced) in $Erβ^+$ hormone induced BC cell line, and around 38% of these spliced genes claimed to be induced by alternative promoters. The present survey discriminated between several cancer specific alternative splicing genes like *LIFR* a BC metastasis suppressor, *PBX1* a pioneer factor defining aggressive $Erβ^-$ BC and *PHLPP2* a tumor suppressor, as exhibiting significant exon modulation in early AS occurrence in hormone responsive $Erβ^+$ BC exclusively. Although, our findings supported dissimilar reply comparing both Cuflinks/Cuffdiff and DEXseq approaches called AS events, it is noteworthy to underline their relative agreement, evaluating spliced genes functional annotation as well as their complementarity performing whole AS survey. This study therefore proposed the integration between Cufflinks/Cuffdiff and DEXseq tools as a reasonable complementary methodology assessing full AS pattern in hormone responsive Erβ BC cells.

Key words: Cufflinks/Cuffdiff, DEXseq, RNA-Seq, alternative splicing (AS), exons, transcript isoforms, estrogen receptor β (ERβ), breast cancer (BC) cells.

INTRODUCTION

Alternative splicing is a central cellular process that produces different mRNA transcript isoforms from a single gene. The qualitative and quantitative identification of such transcript isoforms is more complex as well as essential for understanding the different roles of alternatively spliced genes occurrence in a cell. However, detection of disease-specific transcript isoforms is an important task because aberrant splicing is known to be responsible for various diseases (Kim et al., 2008) and associated with different cancer types (Christofk et al., 2008; Venables et al., 2009). Several studies provided an intriguing insight into the mechanism of cancer specific alternative transcription and alternative splicing, which have long been implicated in the development of cancer. Cancer-specific isoforms have enhanced proliferative, invasive, and migratory abilities and provide a survival advantage to the tumor cells, suggesting that there is specific manipulation of the alternative event regulation in cancer that is beyond the general lack of fidelity of the splicing or the alternative transcription regulatory machinery. It is conceivable that the balanced expression of isoforms, rather than just activation or inhibition of those genes, may hold the key to impeding tumor growth and accordingly it is important to target the disease associated genes at the isoform level rather than at the gene level. The well-known application of exon arrays (Moller-Levet et al., 2009) and the advent of massive parallel sequencing named Next Generation RNA Sequencing (NGS-RNA-Seq) are allowing whole cancer genomes and transcriptomes to be sequenced with extraordinary speed and accuracy, providing insight into the bewildering complexity of isoform-specific expression in cancer genomes (Cancer Genome Atlas, 2012). Detecting alternative isoform regulation is inherently difficult in RNA-Seq, as sequencer reads are often one or more orders of magnitude shorter than the transcripts themselves. While there are several utilities that attempt to de-convolute read data into isoform abundances, the accuracy and robustness of these methods is difficult to establish (Chandramohan et al., 2013; Zhang et al., 2014). In addition, transcript isoforms expression estimates seem to vary considerably between different tools, and generally depend on the quality and completeness of the transcript assembly (Kanitz et al., 2015; Rehrauer et al., 2013). Existing well-established methods detect alternative splicing process mainly by considering sequencing reads that map uniquely to single isoforms or by assembling transcripts and estimating the most likely isoform abundance levels according to the given sequencing reads (Jeffrey and Zhong, 2011). Short read assemblers have been developed for genome and transcriptome assembly, that is, Velvet (Zerbino and Birney, 2008), Scripture (Guttman et al., 2010), Cufflinks (Trapnell et al., 2010) and others (Jeffrey and Zhong, 2011). In literature, the term alternative splicing has been used to describe both alternative transcription and splicing events. However, the variation in the pattern of intron removal, exon joining, and the addition of a poly-A tail on a single pre-mRNA result in alternatively splice mature mRNAs. These various alternative events have been identified in different cells and tissues by application of RNA sequencing (RNA-Seq) next generation sequencing (NGS) based methods in genome wide studies (Sultan et al., 2008; Wang et al., 2008; Pal et al., 2011). Furthermore, it is estimated that there are 263,772 exons in the human genome and approximately, 22% of these exons participate in alternative splicing phenomena (Pal et al., 2012), suggesting that performance assessment of whole alternative splicing occurrence in a genomic survey needed strong bioinformatics/bio-statistical tools characterizing statistical exon modulation. However, a systematic assessment of transcriptome assemblies is difficult because appropriate quality metrics have not been established yet and require a well-defined gold standard that is difficult to find (Jeffrey and Zhong, 2011). Recently, Anders et al. (2012) published a bioinformatics tool called DEXseq, which process exon differential survey in RNA-Seq genomic and/or transcriptomic experimentation. Rather than using an assembly approach and comparing abundance levels of predicted transcripts, DEXseq avoids the assembly step and calculates probabilities values (p-values) for every annotated exon (statistical estimation of exons abundance). Then, the emergence of NGS provides an exciting new technology to analyse alternative splicing on a large scale including differential expression analysis of exons of multi-exon genes. Hence, we believe that combining transcript isoforms differential expression survey with gene and/or transcript isoform exons modulation, could increase researcher alternative splicing phenomena comprehension processing two or more analysed cells and/or tissues conditions. Here we investigated the relation-ship between significantly differentially spliced genes (DSGs) based on alternative transcript isoforms measurement of our previous developed Erβ^{+} breast cancer (BC) cell line model and their corresponding exons abundance modulation merging both Cufflikns/Cuffdiff (Trapnell et al., 2013) and DEXseq (Anders et al., 2012) bioinformatics/bio-statistical approaches respectively, emphasizing alternative splicing manifestation in developed Erβ^{+} and Erβ^{-} breast cancer cell line models (Grober et al., 2011; Paris et al., 2012;

*Corresponding author. E-mail: dgnoel7@gmail.com Tel: (+39) 3381426596 or (+225) 48397811 or (+225) 44595981.

Nassa et al., 2011; Tarallo et al., 2011).

MATERIALS AND METHODS

Erβ$^+$ breast cancer cell model preparation and gene and transcript isoforms differential expression analysis

MCF7 and 5B12 breast cancer (BC) cell line model have been prepared to mimic Erβ/Erβ; (Erβ$^+$) and Erβ/Erα; (Erβ$^-$) breast cancer (BC). The preparation of Erβ$^+$ and Erβ$^-$ BC cell lines have been entirely described in Nassa et al. (2011), Grober et al. (2011) and Dago et al. (2015). Differential genes, transcript isoforms and exons expression analysis, assessing the effect of early stimulation of estradiol (E_2) hormone on alternative splicing occurrence in breast tumor cell models, have been performed by Cufflinks/Cuffdiff (Trapnell et al., 2013) and DEXseq (Anders et al., 2012) bioinformatics/bio-statistic packages respectively. Next generation RNA sequencing (NGS RNA-Seq) data used for the present analysis have been deposited in the Gene Expression Omnibus genomics data public repository (http://www. ncbi.nlm.nih.gov/geo/) with Accession Number GSE64590.

RNA sequencing (RNA-Seq) data generation and gene and/or transcript isoforms expression measurement

1 µg micrograms of high-quality total mRNA was used as starting material for the Illumina mRNASeq library preparation kit and was prepared to manufacturer's directions (Illumina). Libraries were sequenced on the Illumina Genome Analyzer II as 101 base pair paired-end reads. Tophat v.2.0.8 (Kim et al., 2013) was used to align all reads including junction-spanning reads back to the human genome (*Homo sapiens* Ensembl GRCh37). The reads alignment quality and distribution were estimated using SAMtools. Cufflinks v2.1.1 (Trapnell et al., 2013) was used to identify differential spliced genes, isoform transcript and gene expression changes between analyzed experimental groups and/or conditions. We defined statistical significance in expression q-value and/or adjusted p-value for multiple testing ≤ 0.05, and Fragments per Kilobase of exon per Million reads mapped (FPKM)>0.5, since FPKM is a measure of expression used in high-throughput sequencing data that is normalized for both transcript length and total number of reads sequenced.

DEXseq measuring exons differential modulation usage by RNA-sequence

DEXseq is a package for the statistical programming language R (R Development Core Team, 2009) available as open source software via the Bioconductor project (Gentleman et al., 2004). For the preparation steps, namely the flattening of the transcriptome annotation to counting bins and the counting of the reads overlapping each counting bin, two Python scripts are provided, which are built on the HTSeq framework (Anders et al., 2015). The first script takes a GTF file with gene models and transforms it into a GFF file listing counting bins, the second takes such a GFF file and an alignment file in the SAM format and produces a list of counts. The R package is used to read these counts, estimate the size factors and dispersions, fit the dispersion-mean relation and test for differential exon usage. Then, exons with significant change and/or modulation at a false discovery rate lower or equal to 0.05 (FDR≤0.05) have been selected as involve in alternative splice events. All bioinformatics and biostatistics analyses and comparisons were implemented and performed using in-house scripts written in Unix and R.

Functional annotation gene ontology analysis by DAVID

We performed Database for Annotation, Visualization and Integrated Discovery (DAVID http://david.abcc.ncifcrf.gov/) analysis (Glynn et al., 2003) focusing exclusively on the set of differentially spliced gene transcripts and/or isoforms, discriminated by both Cufflinks/Cuffdiff and DEXseq approaches. Then, we performed a gene ontology (GO) survey (FDR ≤0.05 with at least 2 fold enrichment) by processing significantly differentially spliced genes, evaluating alternative transcript and exon change and/or modulation in multi-exon genes.

RESULTS

Reads sequences from both Erβ+ and Erβ- human BC cell line models high-throughput RNA sequencing (RNA-Seq) statistical analysis

mRNA sequencing experiment basing on illumina genome analyzer II (GAII) processing both Erβ+ and Erβ- hormone induced breast cancer (BC) cell lines, yielded approximately 94911602-99876796 million pair-end read (101 bp) sequences (Table 1). From these reads, low quality sequences, were eliminated, resulting in around 65-71 million reads corresponding to 68-70.2% of total reads for each processed replicate sample (Table 1). In total around 65-71 million reads were aligned to the *Homo sapiens* Ensembl GRCh37 reference genome (Table 1). The number of reads per genes and transcript isoforms were further normalized to Fragment per Kilobases of exon per Million mapped reads (FPKM). Then, in order to include a maximum number of genes and transcript isoforms, reducing as possible statistical type I error in calling differentially modulated and spliced genes as well as exon modulation events, we adopted 0.5 FPKM, as genes and/or transcript isoforms expression level threshold in both processed Erβ$^+$ and Erβ$^-$ BC cell exemplars (Figure 1). Student test analysis, based on both loaded junctions and found junction parameters from transcriptome and/or genome reconstruction through Cufflinks package, suggested a reasonable difference (p-value <0.01) between Erβ$^+$ BC cell line under estradiol stimulus (Erβ$^+E_2$) and the other's analyzed BC cell conditions (Table 1). In other words, the present result supported that Erβ$^+$ and Erβ$^-$ BC cell lines in normal conditions (Erβ$^+$noE_2, Erβ$^-$noE_2) as well as Erβ$^-$ BC cell line induced by E_2 (Erβ$^-E_2$), exhibited the same behaviors as opposed to Erβ$^+E_2$ breast cancer cell reacting to E_2 induction (Table 1).

Since each analyzed condition have been processed in three technical replicates, Erβ$^+E_2$_Rep1 indicates the replicate 1 of Erβ$^+$ BC cell line condition under estradiol (E_2) treatment, while Erβ$^+$noE_2_Rep1 indicates replicate 1 of the same cell line with any E_2 treatment The same nomenclature has been adopted for Erβ$^-E_2$ and Erβ$^+$noE_2 analyzed samples conditions.

Table 1. Summary of RNA-Seq reads sequences for each analyzed experimental condition ($Erβ^+E_2$, $Erβ^+noE_2$, $Erβ^-E_2$ and $Erβ^-noE_2$ conditions).

Samples	Sequenced fragment	Aligned fragment	Percent aligned (%)	Loaded junctions	Found junctions
$Erβ^+E_2$_Rep1	97634516	68420165	70.21	197356	165750
$Erβ^+E_2$_Rep2	98586566	68906448	70.02	198164	166736
$Erβ^+E_2$_Rep3	96977352	71363401	70.10	197142	165686
$Erβ^+noE_2$_Rep1	95754654	65171906	68.15	229834	190798
$Erβ^+noE_2$_Rep2	96703300	65626636	68.00	244701	198966
$Erβ^+noE_2$_Rep3	94911602	64472829	68.08	243329	198181
$Erβ^-E_2$_Rep1	97994626	69120928	70.67	243336	198938
$Erβ^-E_2$_Rep2	99083778	69790775	70.56	243590	199088
$Erβ^-E_2$_Rep3	97277796	68582921	70.62	243082	198391
$Erβ^-noE_2$_Rep1	98893200	67641680	68.44	242723	197682
$Erβ^-noE_2$_Rep2	99876796	68208308	68.41	243016	197921
$Erβ^-noE_2$_Rep3	98117332	67079584	68.48	241857	196947

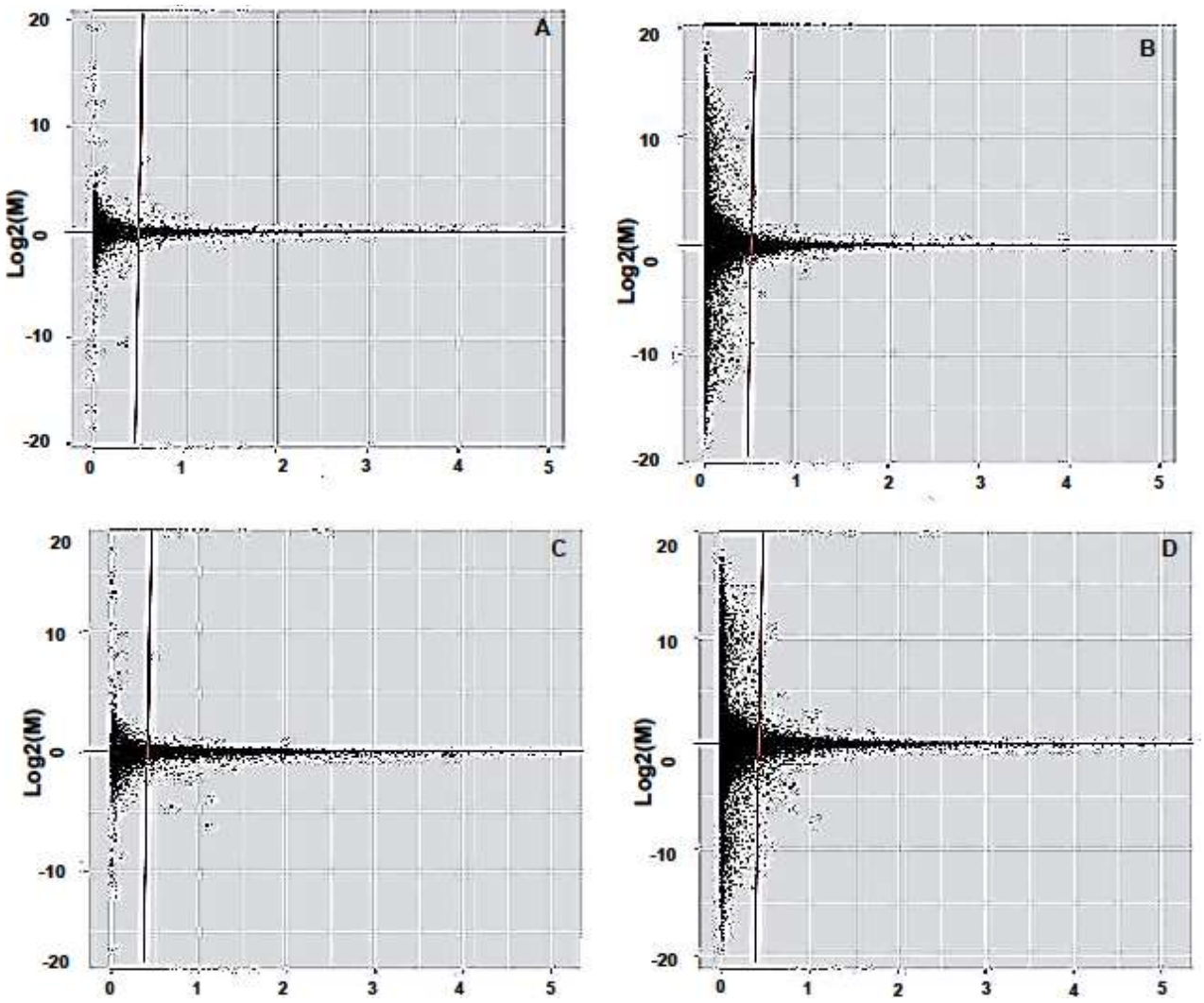

Figure 1. Panels A, B, C and D (MAplot graphics) indicate log 2 mensuration (normalized expression value) referred to genes and/or gene transcript isoforms expression of processed hormone responsive $Erβ^-$ and $Erβ^+$ breast cancer (BC) cell line models respectively.

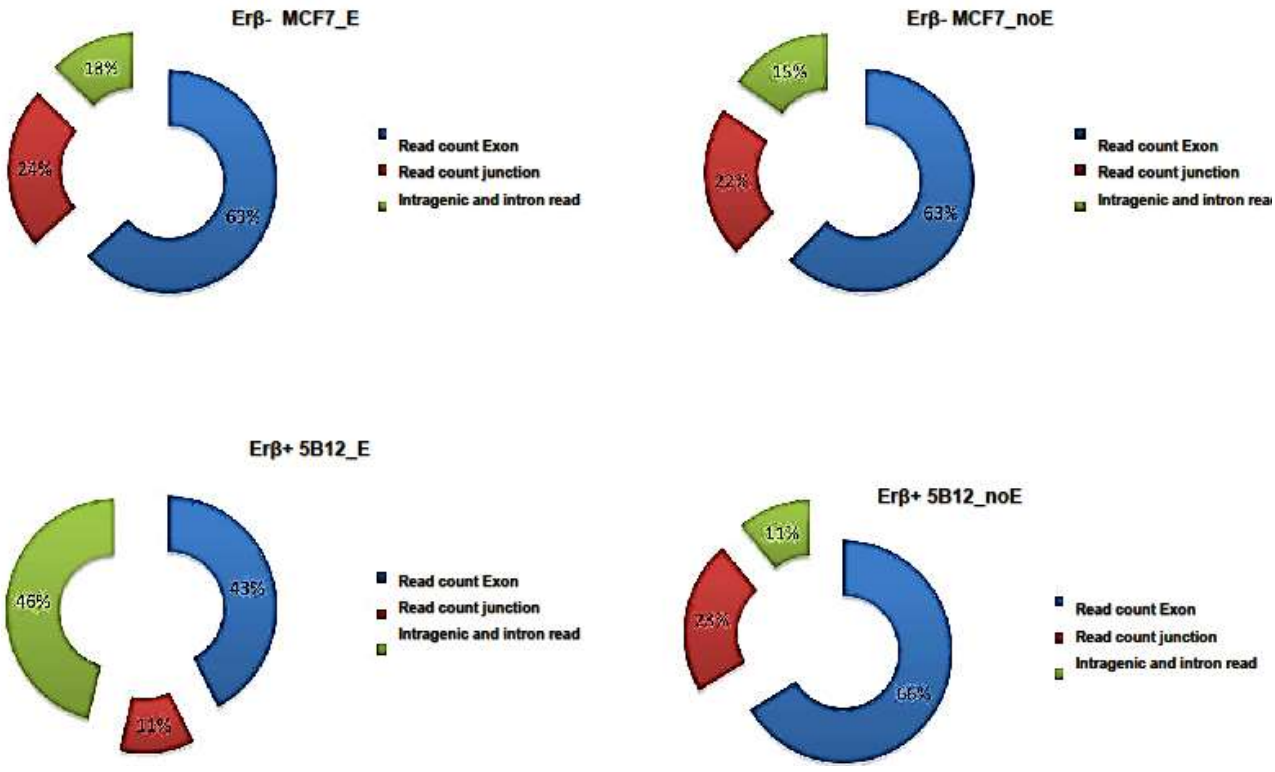

Figure 2. Summary of RNA-Seq reads sequences re-distribution on *Homo sapiens* Ensembl GRCh37 genome assessing alternative splicing occurrence in hormone responsive BC cells by Tophat/Cufflinks and DEXseq/HTseq tools.

Genomic re-distribution of RNA sequencing (RNA-Seq) reads sequences evaluating breast cancer cell line reply to estradiol (E_2) hormone stimulus

Next, we focused on genomic re-distribution of reads sequences, merging read count analysis results from HTSeq/DEXseq with those from Tophat2/Cufflinks software's for each analyzed hormone induced Erβ BC cell line conditions ($Erβ^+E_2$, $Erβ^+noE_2$, $Erβ^-E_2$ and $Erβ^-noE_2$) appraising early alternative splicing events monitoring differential spliced genes and exons change in breast tumor. The present survey highlighted a strong difference in read distributions between $Erβ^+E_2$ BC cells and the other's analyzed conditions ($Erβ^+noE_2$, $Erβ^-E_2$ and $Erβ^-noE_2$) (Figure 2). These results seem to be in agreement with above reported loaded and/or found junction analysis (Table 1), supporting subtly an agreement between HTSeq/DEXseq and Tophat2/ Cufflinks approaches, in genome reconstruction process and/or in genomic reads re-distribution analysis. Furthermore, the present survey exhibited a constancy rate, measuring exon read count as well as intragenic and intron reads values comparing $Erβ^+noE_2$, and both $Erβ^-E_2$ and $Erβ^-noE_2$ breast cancer cells (Figure 2), hypothesizing a similitude between the former's assessing alternative splicing pathway in the present analyzed BC cells, reinforcing the link between estrogen receptor beta (Erβ) and early transcription and mRNA splicing events in hormone responsive BC cells. Considering as a whole, the present results supported a dynamic molecular reply of $Erβ^+$ BC cells as opposed to $Erβ^-$ BC in terms of alternative splicing events (Figure 2).

Assessment of expressed genes and their transcript isoforms proportion comparing hormone responsive $Erβ^+$ and $Erβ^-$ breast cancer cell lines

We evaluated change in differential expressed gene and transcript isoforms including exclusively genes and transcript isoforms that exhibit a FPKM expression value ≥0.5 and significant statistical differential change at an adjusted p-value ≤0.05 by Cufflinks/Cuddiff approach in processed hormone responsive BC cell lines. Then, basing on previous analysis (Figure 1), we processed in total 6714 and 52499 genes and transcript/isoforms respectively (Figure 3). As expected, change in gene and transcript isoforms strongly contrast between both analyzed estrogen hormone responsive $Erβ^+$ and $Erβ^-$ BC cell lines (Figure 3A and B) since, estimated Pearson correlations values comparing log 2 fold change assessing measurement between expressed genes and transcript isoforms resulted to lower 0.5 (R^2 =0.23 and R^2=0.31 for $Erβ^+$ and $Erβ^-$ BC cell respectively), advising

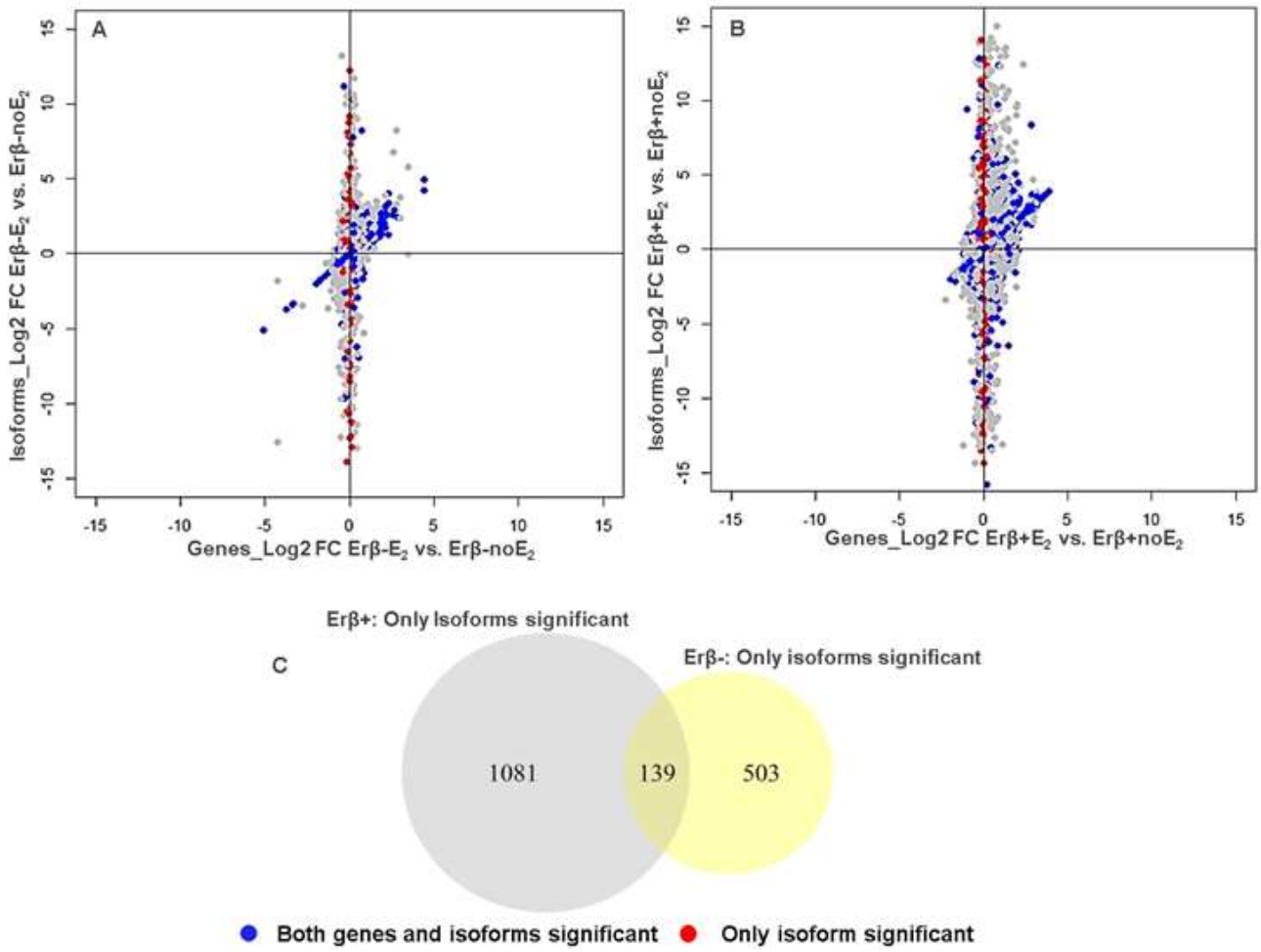

Figure 3. Log_2 fold change scatterplot comparing genes and transcript/isoforms expression in both hormone induced; $Erβ^-$ (A) and $Erβ^+$ or (B) BC cells. The x- coordinate is the gene expression fold change relative to $Erβ^-$ (A) or $Erβ^+$ (B) BC cells and the y-coordinate is the transcript isoform fold change relative to the same BC cells respectively. Blue dots represent the situation where both isoforms and genes, with an expression value FPKM ≥ 0.5, have been selected as statistically significantly differentially expressed at a p-adjusted value≤0.05. Red dots represent situation where only isoform expression is significantly altered. (C) Venn diagrams showing overlapping/divergence between hormone induced Erβ- and Erβ+ BC models analyzing cases where only transcript isoforms expression is significantly altered.

weak agreement between change in genes and their respective transcript isoforms, especially in $Erβ^+$ BC cells (Figure 3). Hence, we focused on the cases where only transcript isoforms expression were significantly altered, alerting alternative promoter usage in alternative transcript isoforms modulation as well as in alternative splicing event occurrence in estrogen responsive BC cell line. Indeed, we showed that situations for which transcript isoform expression was significantly alerted as opposed to their corresponding genes resulted 2 fold more in hormone induced $Erβ^+$ BC cell (Figure 3C). Taking together, the present results suspected a considerable involvement of alternative promoter usage in early alternative splicing occurrence in hormone responsive $Erβ^+$ BC cell lines as opposed to $Erβ^-$ BC cell lines (Figure 3).

Identification of differentially spliced genes (DSGs) in hormone responsive $Erβ^+$ breast cancer cell lines by Cufflinks/Cuffdiff approach

To identify the differences in splice ratios between $Erβ^+E_2$ and $Erβ^+noE_2$ BC cell line ($Erβ^+E_2$ vs. $Erβ^+noE_2$), we employed Cufflinks/Cuffdiff v2.1.1 package, which calculates the changes in the relative splice abundances by quantifying the square root of the Jensen Shannon divergence on all the primary transcripts that produce two or more isoforms. It is essential to note that the distributions of genes, and primary transcripts, and isoform expression level (FPKM) are comparable between the samples that are taken for the differential splicing test. Then, 213 genes randomly distributed on human chromosomes have been detected as significantly

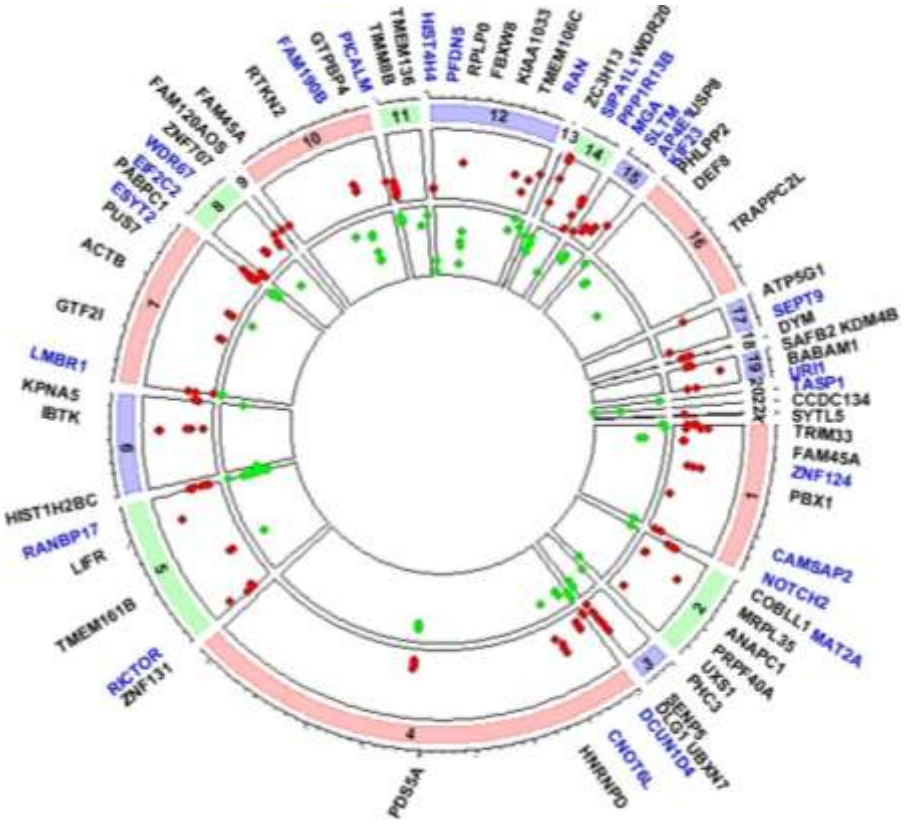

Figure 4. Exons modulation of differentially spliced genes discriminated by both Cufflinks/Cuffdiff and DEXseq in hormone responsive Erβ$^+$ BC cells. Red dots indicate up regulated exons while green dots designate down regulated exons in Erβ$^+$E$_2$ vs. Erβ$^+$noE$_2$ contrast assessing early AS pattern in hormone responsive BC cells. Genes in blue represent spliced multi-exonic genes under alternative promoter usage.

differentially spliced (DSGs) at a false discovery rate ≤0.05 (FDR ≤0.05) monitoring early AS occurrence in hormone responsive Erβ$^+$ BC cell lines (Figure 4 and Supplementary Table 1). Furthermore, basing on this result as well as on previous one (data non shown), we showed that Erβ$^+$ BC cell lines exhibited 3 fold more differentially spliced genes with respect to Erβ- BC cells replying to estradiol hormone stimulus (Supplementary Tables 1 and 2) confirming previous results and observations (Figures 2 and 3). Therefore, we were interested in investigating DSGs exons change monitoring early AS pathway in hormone responsive Erβ+ BC, since exons modulation have been recognized as suitable process understating mechanism of transcript isoforms expression, proposing the former's as potential tumors biomarkers as well as therapeutic target.

Assessment of differential spliced multi-exonic genes (DSGs), exons modulation by DEXseq in Erβ breast cancer cell lines

Here, we analyzed through the DEXseq approach, differential exons expression change, referring to previous detected differential spliced multi-exonic genes (Cufflinks/Cuffdiff) in hormone induced Erβ$^+$ BC cell line. This analysis looks for difference across conditions (in Erβ$^+$E$_2$ vs. in Erβ$^+$noE$_2$) between quantities that are directly observable from shotgun data (read count data), such as the relative usage of each exon. Then, performing the test for differential exons used considering exon with at least 10 reads count in at least one analyzed condition and controlling false discovery rate (FDR) with the Benjamini Hochberg method, at 5% threshold (statistical stringency), 3682 and 122 exons from 1438 and 75 multi-exonic genes, claimed to be significantly differentially modulated in hormone responsive Erβ$^+$ and Erβ$^-$ BC cell lines respectively (Supplementary Materials 1 and 2). Also, this result suggested that exons change ratio in analyzed differentially spliced multi-exonic genes was 1.6 fold more higher in hormone responsive Erβ$^+$BC cell lines when compared to Erβ$^-$ BC cell line, supporting a potential high number of significantly alternatively modulated transcript isoforms involvement characterizing processed hormone responsive Erβ$^+$BC cell line, highlighting the links between Erβ and early AS

Table 2. Assessment of spliced exon position (exon gene ID) in differentially spliced multi-exonic genes called by DEXseq approach.

Gene	DSGs Exon change Exons Gene ID E000-E010	DSGs Exons Change Exon Gene ID E011-E20	DSGs Exons Change Exon Gene ID E021-E30	DSGs Exons Change Exon Gene ID E031-E40	DSGs Exons Change Exon Gene ID E041-E50	DSGs Exons Change Exon Gene ID > E050
(Erβ/Erβ): Erβ$^+$ BC Cell Line	28.04%	22.4%	16.61%	10.77%	6.06%	16.12%
(Erβ/Erα): Erβ$^-$ BC Cell Line	25.62%	18.18%	19%	13.22%	8.26%	15.72%

occurrence in E_2 hormone induced breast cancer. However, significantly modulated exons proportion survey by processing differentially spliced multi-exonic genes revealed that exons change in the analyzed hormone responsive BC cell lines, meanly regard the first 10 gene transcripts exons (more than 25%), advising and/or suspecting weakly solicitation of ending transcripts exons yielding alternative transcript isoforms in hormone responsive BC cells (Table 2).

DAVID analysis assessing functional annotation of differentially spliced genes discriminated by both Cufflinks/Cuffdiff and DEXseq tools

68 genes out the 73 processed differentially spliced genes (DSGs) discriminated by Cufflinks/ Cuffdiff approach at 5% false discovery rate in hormone responsive Erβ$^+$ BC (Supplementary Table 2) were converted in a new list for functional annotation analysis by DAVID package showing at Benjamini correction referring to a p-adjusted value at 5%, that (i) 75% of differentially alternatively spliced genes called by Cufflinks/ Cuffdiff were genes that code for at least two isoforms due to pre-mRNA splicing event and that (ii) 74% of these genes were recognized as alternative splice variant. Furthermore, 42.64% of alternative spliced genes categorized by Cufflinks/ Cuffdiff approach have been discriminated to be localized in the nucleus (Figure 5A). In the same tendency and focusing on alternative spliced exons events called by DEXseq methodology at 5% false discovery rate threshold with 2 fold exon change (high alternative splicing evidence), we showed that 55.35% of analyzed alternative spliced genes, coded for at least two isoforms due to pre-mRNA splicing event and that 54.91% of their transcript isoform variants were recognized as alternative splice variant. We also showed that 37.84% of all analyzed alternative spliced genes called by DEXseq approach have been discriminated to be localized in the nucleus (Figure 5B). These results, delicately suggested a relative agreement between the two considered and/or analyzed bioinformatics tools (Cufflinks/ Cuffdiff and DEXseq) assessing the performance of alternative splicing pathway in breast cancer cell line, since mRNA maturation in eukaryotic cells happened in cell's nucleus. Finally, even if the present results suggested DEXseq approach as discriminating highest number of potential alternative splicing event in term of functional annotation as opposed to Cufflinks/Cuffdiff (Figure 5), it is noteworthy to underline their similitude evaluating (i) splice and isoform variant, (ii) alternative spliced genes localization in the nucleus and (iii) nucleolus in the present functional annotation survey (Figure 5).

Integration between Cufflinks/Cuffdiff and DEXseq data assessing alternative splicing event in Erβ breast cancer cell line

More than 34.74% of DSGs (multi-exonic genes) discriminated by Cufflinks/Cuffdiff approach in the hormone responsive Erβ$^+$ BC cell, exhibited at least one significant exons change as DEXseq exon differential analysis (Figure 4). Indeed, *LIFR* gene, knows as a breast cancer metastasis suppressor, *MATA2* gene involves in apoptosis process in human hepatoma, *NBPF10* gene associated with several types of cancer, *PBX1* gene that results engaged in the progression of breast cancer and *PHLPP2* gene that inhibits cancer cell proliferation acting as tumor suppressor, detected as significantly differentially spliced by Cufflinks/Cuffdiff approach, displayed a significant exons change in the present analyzed hormone responsive BC cell line (Figure 4), proposing both Cufflinks/Cuffdiff and DEXseq data merging process as a valid methodology monitoring AS pathway in breast cancer. As DEXseq provides exons splicing visualization, we reported in Figure 6 an example of exons modulation evaluating alternative splicing happening in two fully differentially spliced genes *MATA2* and *PBX1* in E_2 hormone induced BC exclusively (Erβ$^+$ BC). Moreover, several generic tumor biomarkers such as *SEPT9*, known as a candidate for the ovarian tumor suppressor, *PPP1R13B* and *NOTCH2* genes affecting cells differentiation implementation, proliferation and apoptotic programs and *NBPF10* gene which results associated with several type of cancer, were recognized as exhibiting significant exons change in the hormone responsive Erβ$^+$ BC cell line (Figure 4). We also showed that 36.11% of detected spliced genes merging both Cuffdiff and

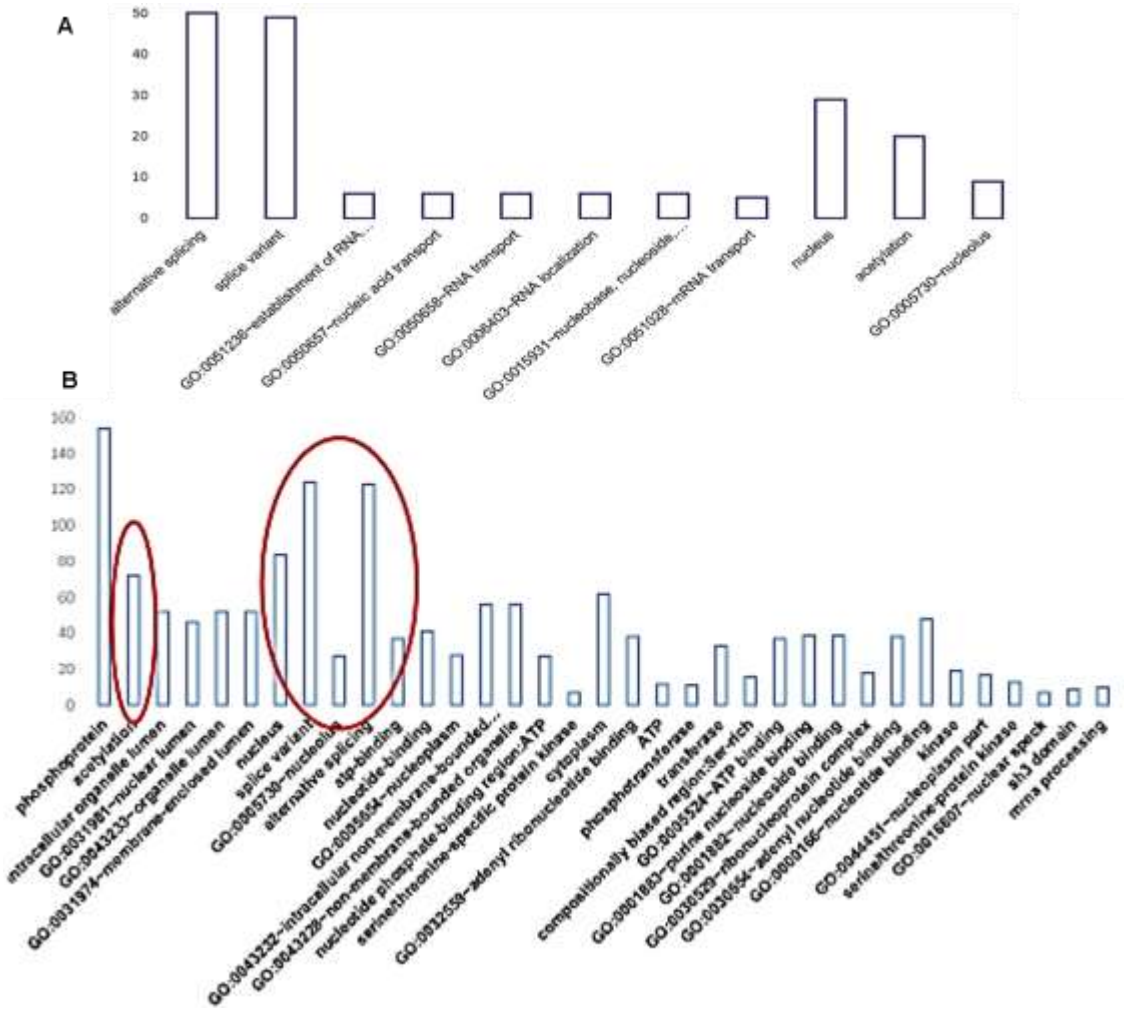

Figure 5. Functional annotation survey by DAVID tool by processing differentially spliced genes (DSGs) candidates comparing both Cufflinks/Cuffdiff (A) and DEXseq (B) bioinformatics approaches.

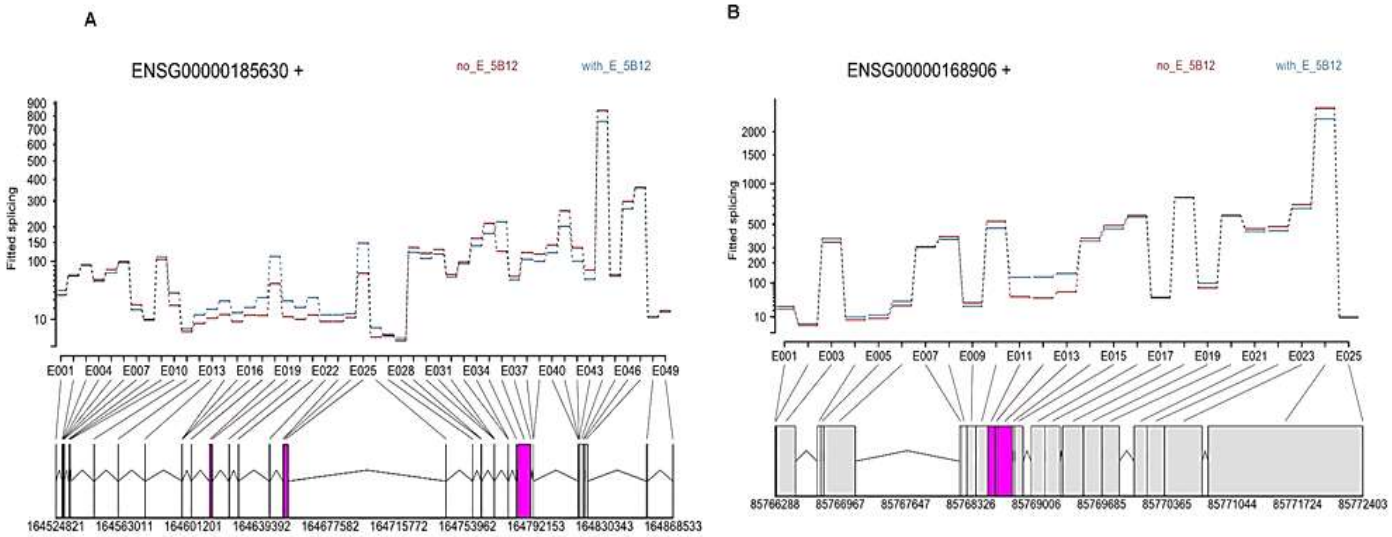

Figure 6. DEXseq representation of significant exon change of two differentially spliced genes PBX1 (A) and MATA2 (B) discriminated by Cufflinks/Cuffdiff and DEXseq. Shown in pink is the differential expressed exons (exon reads count ≥10 at a FDR ≤ 0.05) involved in the potential alternative splicing events.

DEXseq approaches were regulated by significant alternative promoter use (Figure 4). Interestingly, well-noted biomarker genes like *MATA2*, *SEPT9*, *NOTCH2* and *PPP1R13B* claimed to be under alternative promoter usage (p-value≤0.05) (Figure 4 and Supplementary Table 3), assessing alternative splicing pathway in Erβ^{+} BC cells. So, the present results suspecting the involvement of some remarkable cancer biomarkers controlling AS path in hormone responsive BC cell lines, proposed Cufflinks/Cuffdiff and DEXseq approaches data integration as reasonable and complementary scheme weighing whole AS pathways in breast cancer cells.

DISCUSSION

Alternative splicing (AS) is a means of expressing several or many different transcripts from the same genomic DNA and results from the inclusion of a subset of the available exons for a particular protein. By excluding one or more exons, certain protein domains may be lost from the encoded protein, which can result in protein function damage or gain. Several types of alternative splicing have been described resulting in exon skipping; alternative 5′ or 3′ splice sites; mutually exclusive exons and much more rarely and intron retention, approving the complexity of AS study and the needed of accurate bioinformatics or bio-statistics systems evaluating this phenomena in eukaryotic cells. However, many alternative splicing events have been noted in human development, especially in the brain and the testes (Grabowski and Black, 2001; Venables, 2002) as well as in cancer, including the use of alternative individual splice sites, alternative exons, and alternative introns. Therefore, whole alternative splicing survey must include capable sensitive and specific bioinformatics tools, able to measure accurately alternative transcript isoforms as well as multi-exonic genes transcripts exons abundance, since recent Next Generation RNA Sequencing (NGS RNA-Seq) analysis provides innovative platform exploring in detail cell transcriptome and/or genome. However, a recurrent challenge in RNA-Seq experimentation regards application of adequate bioinformatics tools analyzing huge quantity and complex yielded data. Also, RNA-Seq coupled with well-established bioinformatics approaches; that is, Cufflinks/Cuffdiff (Trapnell et al., 2013) and DEXseq (Anders et al., 2012), allowed a suitable whole genome and transcriptome reconstruction measuring gene transcript isoforms as well as exons expression level, performing differential expression analysis between cells types. In such analysis, annotated regions of the considered genome can be expressed (that is, exons), describing how the pre-mRNAs are spliced into transcripts. While there are several utilities that attempt to de-convolute read data into isoform abundances, the accuracy and robustness of these methods is difficult to establish (Chandramohan et al., 2013; Zhang et al., 2014). Isoform expression estimates seem to vary considerably between different tools, and generally depend on the quality and completeness of the transcript assembly (Kanitz et al., 2015; Rehrauer et al., 2013). Also, Cufflinks/Cuffdiff methodology is not more informative regarding exons modulation measuring alternative splice transcript isoforms pattern. Simon Anders et al. (2012) demonstrate the versatility of DEXseq package by applying it to several data sets facilitating the study of regulation and function of alternative exon usage on a genome-wide scale. Guided by these observations, we proposed merging between both Cufflinks/Cuffdiff and DEXseq investigating meticulously alternative transcript isoforms exons abundance in our previous studied hormone responsive Erβ+ breast cancer cells, since Erβ significantly affects estrogen-induced early transcription and mRNA splicing in hormone-responsive BC cells (Dago et al., 2015). Then, genomic re-distribution of RNA-Seq reads sequences from hormone responsive Erβ BC cells by HTSeq/DEXseq and TopHat/Cufflinks, suggested strong difference (p-value <0.05) between estrogen induced Erβ^{+} BC cells (Erβ$^{+}E_2$) and the other's analyzed breast cancer cells conditions (Erβ^{+} with any E_2 stimulus; Erβ^{+}noE_2, Erβ^{-} induced by E_2; Erβ$^{-}E_2$ and Erβ^{-} without any E_2 stimulus; Erβ^{-}noE_2), since exhibiting dissimilar attitude and/or behaviors considering loaded and fund reads junction as well as exon intragenic and intron reads distribution (Figure 2 and Table 1), subtly evoking agreement between both Cufflinks and DEXseq approaches in genomic and transcriptomic reconstruction analysis. In addition, this analysis suggested high variability between Erβ$^{+}E_2$ and Erβ^{+}noE_2 conditions suspecting a high number of differentially expressed genes and/or transcript isoforms in the latter's as opposed to hormone induced Erβ^{-} BC cells. However, comparative analysis assessing the relationship between change in expressed genes and their respective transcript isoforms, showed a strong evidence of alternative promoter used in alternative splicing pattern processing hormone responsive Erβ^{+} BC cell line, as the ratio between expressed transcript isoforms as oppose to expressed genes, was 2 fold more higher in the latter's (Erβ^{+} BC) when compared to Erβ^{-}BC cell line (Figure 3). Taking together, these results reinforced alternative splicing evidence in hormone responsive Erβ^{+}BC as opposite to Erβ^{-} BC cell lines alerting a significant participation of alternative promoter regulating AS events occurrence in breast cancer process. Several studies have shown that the occurrence of alternative transcriptional termination and splicing is higher in genes with alternative promoters, and the choice of alternative promoter and transcription termination can influence the alternative splicing pattern of the pre-mRNA (Winter et al., 2007; Albulescu et al., 2012). Nevertheless, it is estimated that participate to the alternative splicing phenomena proposing genes transcript isoforms exons modulation as a suitable approach understanding

alternative transcript isoform expression in human genome. Hence, detailed alternative splicing analysis by NGS RNA-Seq need innovative bioinformatics scheme capable to integrate accurately transcript isoforms and exon expression analysis. Based on this, we combined differentially spliced genes and transcript isoforms exons modulation assessing early AS occurrence in estradiol hormone induced Erβ+ BC cells, emphasizing strong involvement of transcript isoforms exons change in alternative splicing mechanism in breast tumor. Then, integrating Cufflinks/ Cuffdiff and DEXseq approaches, our findings proposed alternative promoter's usage as well as significant exon change as key molecular events favoring AS pattern in hormone responsive breast cancer cells (Figure 4). Also, it is noteworthy to underline the weak involvement of ended exons of transcript isoforms evaluating alternative splicing occurrence in the present hormone induced Erβ BC cells (Table 2). Furthermore, around 34.74% of significantly differentially spliced genes by Cufflinks/ Cuffdiff have shown significant exons change in DEXseq analysis testing for differential usage of exon regions as a proxy for alternative isoform regulation as well as providing a powerful suite of visualization tools (Figure 6) (Love et al., 2014). Interestingly, the present analysis identified several remarkable cancer biomarkers, like *MDM2* and *NF1* genes, known as cancer specific alternative splicing genes (Venables, 2004), *PBX1* gene, revealed as a novel pioneer factor defining aggressive $Er\beta^-$ breast tumors, as it guides Erα genomic activity to unique genomic regions promoting a transcriptional program favorable to breast cancer progression (Magnani et al., 2011), *LIFR* gene, knows as a breast cancer metastasis suppressor (Chen et al., 2012), *MATA2* gene, involved in human colon cancer progression (Chen et al., 2007) and *PHLPP2* tumor suppressor (Liu et al., 2011) and *AIB1/NCOA3* hormone signaling in breast cancer as exhibiting significant exon change in hormone responsive Erβ BC cells, demonstrating the key role of exon modulation in early estrogen induced breast cancer alternative splicing pattern (Figure 4, Supplementary Materials 1 and 2). Furthermore, 36.11% of detected differentially spliced genes by merging Cuffdiff and DEXseq approaches were recognized as including alternative promoter, and some well-noted cancer biomarker (*MATA2*, *SEPT9*, *NOTCH2* and *PPP1R13B*) claimed to be under alternative promoter usage ($p\text{-value} \leq 0.05$) (Figure 4 and Supplementary Table 3), monitoring early alternative splicing pathway in our processed hormone responsive $Er\beta^+$ BC cells (Figure 4). So, the present results by evoking the involvement of some remarkable cancer biomarkers, controlling AS pattern in the present hormone responsive BC cell lines, proposed Cufflinks/Cuffdiff and DEXseq genomic data integration, as a reasonable complementary scheme assessing whole AS occurrence in hormone responsive breast cancer cell. Moreover, concordance between Cufflinks/ Cuffdiff and DEXseq has been supported in part by DAVID functional annotation analysis since 42.64 and 37.84% of alternative spliced genes categorized by previous mentioned approaches respectively, have been discriminated to be localized in cell's nucleus, suggesting a strong contribution of the latter's (alternative sliced genes) regulating AS pattern occurrence in hormone responsive BC cell line (Figure 5). Also, even if the present study admitted conflicting results between Cufflinks/Cuffdiff and DEXseq approaches in alternative splicing analysis (Figure 5), we showed that an adequate integration between these bioinformatics/biostatistics tools can help to wholly investigate alternative splicing occurrence in breast cancer disease reducing false discovery event rate. Furthermore, accuracy in statistical analysis processing alternative transcript isoforms regulation in genomic and transcriptomic studies, has been supported by recent work introducing a new method that builds on the statistical techniques used by the well-established DEXseq package to detect differential usage of both exonic regions and splice junctions helping differential usage of novel splice junctions without the need for an additional isoform assembly step (Stephen and James, 2016).

Conclusion

RNA-Seq providing more than simple measurements of gene and/or transcript-level expression, can be used to study more complex regulatory phenomena at the isoform level, even when the isoforms in question are unannotated. Numerous tools have been developed to detect alternative isoform regulation exhibiting in some cases conflicting results. In contrast to this tendency, the present study proposed integration between well-established Cufflinks/Cuffdiff and DEXseq approaches as reasonable system assessing alternative splicing events in hormone responsive Erβ breast cancer cells. Finally, our findings exhibited significant exon modulation of multi-exonic gene transcripts regulated by alternative promoters, as recurrently solicited in early AS pattern in estrogen induced BC cells.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

ACKNOWLEDGEMENT

Thanks to all members of the Laboratory of Molecular Medicine and Genomics of the Department of Medicine and Surgery, University of Salerno (SA) Italy, for providing information regarding RNA sequencing data for the present study.

REFERENCES

Albulescu LO, Sabet N, Gudipati M, Stepankiw N, Bergman ZJ, Huffaker TC, Pleiss JA (2012). A quantitative high-throughput reverse genetic screen reveals novel connections between Pre-mRNA splicing and 5′ and 3′ end transcript determinants. PLoS Genet. 8(3):e1002530.

Anders S, Reyes A, Huber W (2012). Detecting differential usage of exons from RNA-seq data. Genome Res. 22(10):2008-2017.

Anders S, Pyl PT, Huber W (2015). HTSeq A Python framework to work with high-throughput sequencing data. Bioinformatics 31(2):166-169.

Cancer Genome Atlas (2012). Comprehensive molecular characterization of human colon and rectal cancer. Nature 487:330-337.

Chandramohan R, Wu PY, Phan H, Wang MD (2013). Benchmarking RNA-Seq quantification tools. Conf. Proc. IEEE Eng. Med. Biol. Soc. pp. 647-650.

Chen D, Sun Y, Wei Y, Zhang P, Rezaeian AH, Teruya-Feldstein J, Gupta S, Liang H, Lin HK, Hung MC, Ma L (2012). LIFR is a breast cancer metastasis suppressor upstream of the Hippo-YAP pathway and a prognostic marker. Nat. Med. 18(10):1511-1517.

Chen H, Xia M, Lin M, Yang H, Kuhlenkamp J, Li T, Sodir NM, Chen YH, Josef-Lenz H, Laird PW, Clarke S, Mato JM, Lu SC (2007). Role of methionine adenosyltransferase 2A and S-adenosylmethionine in mitogen-induced growth of human colon cancer cells. Gastroenterology 133(1):207-218.

Christofk HR, Vander Heiden MG, Harris MH, Ramanathan A, Gerszten RE, Wei R, Fleming MD, Schreiber SL, Cantley LC (2008). The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. Nature 452:230-233.

Dago DN, Claudio S, Antonio R, DomenicoM, Giorgio G, Giovanni N, Maria R, Francesca R, Roberta T, Alessandro W (2015). Estrogen receptor beta impacts hormone-induced alternative mRNA splicing in breast cancer cells. BMC Genomics 16(367):1-13.

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J (2004). Bioconductor: open software development for computational biology and bioinformatics Genome Biol. 5(10):1-39.

Glynn DJ, Brad TS, Douglas AH, Jun Y, Wei G, Clifford HL and Richard AL (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 4:1-11.

Grabowski PJ, Black DL (2001). Alternative RNA splicing in the nervous system. Prog. Neurobiol. 65:289-308.

Grober OM, Mutarelli M, Giurato G, Ravo M, Cicatiello L, De Filippo MR, Ferraro L, Nassa G, Papa MF, Paris O, Tarallo R (2011). Global analysis of estrogen receptor beta binding to breast cancer cell genome reveals an extensive interplay with estrogen receptor alpha for target gene regulation. BMC Genomics 12(36):1-16.

Jeffrey AM, Zhong W (2011). Next-generation transcriptome assembly. Nat. Rev. Genet. 12:671-682.

Kanitz A, Gypas F, Gruber AJ, Gruber AR, Martin G and Zavolan M (2015). Comparative assessment of methods for the computational inference of transcript isoform abundance from RNAseq data. Genome Biol. 16:1-26.

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14(4):1-13.

Kim E, Goren A, Ast G (2008). Insights into the connection between cancer and alternative splicing. Trends Genet. 24:7-10.

Liu J, Stevens PD, Li X, Schmidt MD, Gao T (2011). PHLPP-mediated dephosphorylation of S6K1 inhibits protein translation and cell growth. Mol. Cell. Biol. 31(24):4917-4927.

Love MI, Huber W, Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15(12):550.

Magnani L, Elizabeth BB, Xiaoyang Z, Mathieu L (2011). PBX1 Genomic Pioneer Function Drives ERα Signaling Underlying Progression in Breast Cancer PLoS Genet. 7(11):e1002368.

Moller-Levet CS, Betts GN, Harris AL, Homer JJ, West CM, Miller CJ (2009). Exon array analysis of head and neck cancers identifies a hypoxia related splice variant of LAMA3 associated with a poor prognosis. PLoS Comput. Biol. 5(11):e1000571.

Nassa G, Tarallo R, Ambrosino C, Bamundo A, Ferraro L, Paris O, et al. (2011). A large set of estrogen receptor beta-interacting proteins identified by tandem affinity purification in hormone-responsive human breast cancer cell nuclei. Proteomics 11(1):159-165.

Pal S, Gupta R, Kim H, Wickramasinghe P, Baubet V, Showe LC, et al. (2011). Alternative transcription exceeds alternative splicing in generating the transcriptome diversity of cerebellar development. Genome Res. 21(8):1260-72.

Pal S, Ravi G, Ramana VD (2012). Alternative transcription and alternative splicing in cancer. Pharmacol. Ther. 136:283-294.

Paris O, Ferraro L, Grober OM, Ravo M, De Filippo MR, Giurato G, Nassa G, Tarallo R, Cantarella C, Rizzo F, Di Benedetto A (2012). Direct regulation of microRNA biogenesis and expression by estrogen receptor beta in hormone-responsive breast cancer. Oncogene 31(38):4196-206.

R Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing. URL http://www.R-project.org.

Rehrauer H, Opitz L, Tan G, Sieverling L, Schlapbach R (2013). Blind spots of quantitative RNA-seq: the limits for assessing abundance, differential expression, and isoform switching. BMC Bioinformatics 14:370.

Stephen WH, James CM (2016). Detection and visualization of differential splicing in RNA-Seq data with JunctionSeq. Nucleic Acids Res. 44(15):1-15.

Sultan M, Schulz MH, Richard H, Magen A, Klingenhoff A, Scherf M, Seifert M, Borodina T, Soldatov A, Parkhomchuk D, Schmidt D (2008). A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. Science 321:956-960.

Tarallo R, Bamundo A, Nassa G, Nola E, Paris O, Ambrosino C, Facchiano A, Baumann M, Nyman TA, Weisz A (2011). Identification of proteins associated with ligand-activated estrogen receptor alpha in human breast cancer cell nuclei by tandem affinity purification and nano LC-MS/MS. Proteomics 11(1):172-179.

Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat. Biotechnol. 31(1):46-53.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. 28(5):511-515.

Venables JP (2002). Alternative splicing in the testes. Curr. Opin. Genet. Dev.12:615-619.

Venables PJ (2004). Aberrant and Alternative Splicing in Cancer. Cancer Res. 64:7647-7654.

Venables,JP, Klinck R, Koh C, Gervais-Bird J, Bramard A, Inkel L, Durand M, Couture S, Froehlich U, Lapointe E, Lucier JF, Thibault P, Rancourt C, Tremblay K, Prinos P, Chabot B, Elela SA (2009). Cancer-associated regulation of alternative splicing. Nat. Struct. Mol. Biol.16:670-676.

Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB (2008). Alternative isoform regulation in human tissue transcriptomes. Nature 456:470-476.

Winter J, Kunath M, Roepcke S, Krause S, Schneider R, Schweiger S (2007). Alternative polyadenylation signals and promoters act in concert to control tissue-specific expression of the Opitz syndrome gene MID1. BMC Mol. Biol. 8(105):1-12.

Zerbino DR, Birney E (2008).Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res.18(5):821-829.

Zhang ZH, Jhaveri DJ, Marshall VM, Bauer DC, Edson J, Narayanan RK, Robinson GJ, Lundberg AE, Bartlett PF, Wray NR, Zhao QY (2014). A comparative study of techniques for differential expression analysis on RNA-Seq data. PLoS One 9(8):e103207.

Supplementary Table 1. Erβ+ BC cell line spliced genes at q_value ≤0.05.

Gene_id	Gene	Locus	sqrt(JS)	p_value	q_value	Significant
ENSG00000234608	C12orf47	12:112277570-112334343	0.215053	0.00255	0.0318605	yes
ENSG00000234741	GAS5	1:173831289-173866494	0.138086	0.00005	0.00104714	yes
ENSG00000068745	IP6K2	3:48725435-48777786	0.27223	0.00075	0.0116144	yes
ENSG00000119285	HEATR1	1:236681299-236767804	0.0748117	0.00005	0.00104714	yes
ENSG00000169689	STRA13	17:79976578-79980794	0.0627464	0.0001	0.00199909	yes
ENSG00000105750	ZNF85	19:21106058-21133503	0.388189	0.00075	0.0116144	yes
ENSG00000075711	DLG1	3:196769430-197030618	0.384547	0.00005	0.00104714	yes
ENSG00000136051	KIAA1033	12:105501101-105562912	0.0564964	0.0043	0.0446024	yes
ENSG00000091039	OSBPL8	12:76745576-76953589	0.0515044	0.002	0.0261786	yes
ENSG00000224831	RP11-651P23.4.1	3:149478891-149942977	0.39333	0.00005	0.00104714	yes
ENSG00000007392	LUC7L	16:238967-279462	0.133702	0.0013	0.018325	yes
ENSG00000092978	GPATCH2	1:217600333-217804424	0.0213977	0.00365	0.0393449	yes
ENSG00000196712	NF1	17:29421944-29709134	0.210908	0.00005	0.00104714	yes
ENSG00000112414	GPR126	6:142622990-142767403	0.498713	0.00005	0.00104714	yes
ENSG00000138592	USP8	15:50716576-50838905	0.193283	0.00005	0.00104714	yes
ENSG00000154743	TSEN2	3:12525930-12705725	0.337813	0.00005	0.00104714	yes
ENSG00000122008	POLK	5:74664310-74896969	0.370688	0.00005	0.00104714	yes
ENSG00000185722	ANKFY1	17:4066664-4167274	0.166863	0.00005	0.00104714	yes
ENSG00000073350	LLGL2	17:73521782-73571289	0.10957	0.00005	0.00104714	yes
ENSG00000100325	ASCC2	22:30184596-30234271	0.578665	0.00005	0.00104714	yes
ENSG00000147274	RBMX	X:135923089-135962923	0.143233	0.00025	0.004398	yes
ENSG00000218739	AC007390.5.1	2:37394962-37551951	0.167284	0.0027	0.0333556	yes
ENSG00000249846	RP11-77P16.4.1	3:129800673-129838359	0.559348	0.00295	0.0346901	yes
ENSG00000185305	ARL15	5:53179774-53606412	0.311492	0.00295	0.0346901	yes
ENSG00000172748	ZNF596	8:182136-197342	0.258979	0.00005	0.00104714	yes
ENSG00000188266	AGPHD1	15:78799905-78829714	0.616911	0.00005	0.00104714	yes
ENSG00000005700	IBTK	6:82879699-82957471	0.15325	0.00005	0.00104714	yes
ENSG00000137776	SLTM	15:59063390-59389618	0.329689	0.00005	0.00104714	yes
ENSG00000182307	C8orf33	8:146277763-146281416	0.0615689	0.0001	0.00199909	yes
ENSG00000090263	MRPS33	7:140702195-140715028	0.0773085	0.00285	0.0346251	yes
ENSG00000175467	SART1	11:65729159-65747299	0.111518	0.00025	0.004398	yes
ENSG00000153107	ANAPC1	2:112523847-112642267	0.153132	0.00005	0.00104714	yes
ENSG00000088808	PPP1R13B	14:104200088-104313927	0.176983	0.0025	0.0315948	yes
ENSG00000123908	EIF2C2	8:141541263-141645718	0.145366	0.0012	0.0174755	yes
ENSG00000074621	SLC24A1	15:65903703-66184329	0.475611	0.0009	0.013649	yes
ENSG00000163867	ZMYM6	1:35439956-35497569	0.273954	0.00005	0.00104714	yes
ENSG00000197548	ATG7	3:11313994-11770134	0.134414	0.0003	0.00515391	yes
ENSG00000221983	UBA52	19:18682613-18688269	0.0527833	0.0003	0.00515391	yes
ENSG00000143367	TUFT1	1:151512780-151556059	0.191093	0.00005	0.00104714	yes
ENSG00000116191	RALGPS2	1:178694299-179067158	0.0718482	0.0041	0.0433457	yes
ENSG00000167792	NDUFV1	11:67374322-67380006	0.23601	0.00095	0.0143086	yes
ENSG00000198160	MIER1	1:67390577-67454302	0.235012	0.00005	0.00104714	yes
ENSG00000108799	EZH1	17:40852293-40897071	0.208434	0.00005	0.00104714	yes
ENSG00000118939	UCHL3	13:76123618-76434004	0.117807	0.00005	0.00104714	yes
ENSG00000119979	FAM45A	10:120863597-120897496	0.205479	0.00215	0.0278109	yes
ENSG00000162063	CCNF	16:2479394-2509980	0.260364	0.00005	0.00104714	yes
ENSG00000254986	DPP3	11:66234215-66313709	0.458483	0.00005	0.00104714	yes
ENSG00000109184	DCUN1D4	4:52709165-52783003	2.14375	0.00005	0.00104714	yes
ENSG00000119772	DNMT3A	2:25455844-25565459	0.0649625	0.00175	0.0237546	yes
ENSG00000077809	GTF2I	7:74071993-74306729	0.103656	0.00325	0.03665	yes
ENSG00000213762	ZNF134	19:58125618-58134721	0.515367	0.00005	0.00104714	yes

Supplementary Table 1. Contd.

ENSG00000196911	KPNA5	6:117002349-117063029	0.320921	0.00005	0.00104714	yes
ENSG00000164180	TMEM161B	5:87485449-87794514	0.389591	0.00005	0.00104714	yes
ENSG00000049283	EPN3	17:48609903-48633213	0.369319	0.00005	0.00104714	yes
ENSG00000141446	ESCO1	18:19109263-19180845	0.0673011	0.00205	0.0266743	yes
ENSG00000181264	TMEM136	11:120195837-120204391	0.291405	0.00005	0.00104714	yes
ENSG00000070756	PABPC1	8:101698043-101735037	0.0677226	0.00075	0.0116144	yes
ENSG00000008294	SPAG9	17:49039534-49198226	0.163075	0.00005	0.00104714	yes
ENSG00000040199	PHLPP2	16:71657610-71758604	0.244715	0.0029	0.0346901	yes
ENSG00000132313	MRPL35	2:86426579-86440917	0.197046	0.0042	0.0437716	yes
ENSG00000164327	RICTOR	5:38845959-39074510	0.55906	0.00005	0.00104714	yes
ENSG00000105576	TNPO2	19:12810007-12834810	0.226942	0.0007	0.0113184	yes
ENSG00000089123	TASP1	20:13202417-13619587	0.222231	0.00085	0.0129802	yes
ENSG00000010404	IDS	X:148558520-148632055	0.0921044	0.0002	0.003665	yes
ENSG00000067066	SP100	2:231280656-231444721	0.189363	0.0003	0.00515391	yes
ENSG00000047230	CTPS2	X:16606125-16731059	0.123897	0.00275	0.0335958	yes
ENSG00000073921	PICALM	11:85668726-85780924	0.108588	0.00105	0.015601	yes
ENSG00000083896	YTHDC1	4:69176104-69215807	0.025357	0.00305	0.0354865	yes
ENSG00000134291	TMEM106C	12:48357351-48362661	0.0742842	0.00005	0.00104714	yes
ENSG00000107771	FAM190B	10:86088341-86278273	0.0696925	0.00005	0.00104714	yes
ENSG00000107937	GTPBP4	10:1034337-1095110	0.107753	0.00005	0.00104714	yes
ENSG00000139697	SBNO1	12:123773655-123834988	0.150288	0.00005	0.00104714	yes
ENSG00000196290	NIF3L1	2:201754049-201768655	0.136813	0.00005	0.00104714	yes
ENSG00000101333	PLCB4	20:9049409-9461889	0.165373	0.00005	0.00104714	yes
ENSG00000157657	ZNF618	9:116638561-116818871	0.117957	0.00405	0.0430239	yes
ENSG00000123349	PFDN5	12:53689074-53700961	0.16216	0.00185	0.0245069	yes
ENSG00000103365	GGA2	16:23474862-23533316	0.107243	0.00005	0.00104714	yes
ENSG00000167863	ATP5H	17:73032144-73043074	0.0369063	0.00255	0.0318605	yes
ENSG00000164171	ITGA2	5:51971026-52390609	0.212906	0.0019	0.0250186	yes
ENSG00000156787	WDR67	8:124014799-124164393	0.160779	0.0042	0.0437716	yes
ENSG00000101596	SMCHD1	18:2655885-2805015	0.0751137	0.00005	0.00104714	yes
ENSG00000117143	UAP1	1:162531320-162569627	0.117689	0.0037	0.0396893	yes
ENSG00000196504	PRPF40A	2:153508106-153617688	0.104612	0.00005	0.00104714	yes
ENSG00000100147	CCDC134	22:42196682-42222303	0.328042	0.00005	0.00104714	yes
ENSG00000109184	DCUN1D4	4:52709165-52783003	0.539569	0.00005	0.00104714	yes
ENSG00000173145	NOC3L	10:95753745-96122716	0.0601538	0.00085	0.0129802	yes
ENSG00000167515	TRAPPC2L	16:88880141-88933068	0.0375659	0.00295	0.0346901	yes

Supplementary Table 2. Erβ- BC Cell line Spliced genes at q_value ≤0.05.

Gene_id	Gene	Locus	sqrt(JS)	Test_stat	p_value	q_value	Significant
ENSG00000234741	GAS5	1:173831289-173866494	0.136996	0	0.00105	0.0398523	yes
	HEATR1	1:236681299-236767804	0.121115	0	5,00E-05	0.00305488	yes
ENSG00000224831	RP11-651P23.4.1	3:149478891-149942977	0.418664	0	5,00E-05	0.00305488	yes
ENSG00000131263	RLIM	X:73805051-73834452	0.065689	0	0.00115	0.04175	yes
ENSG00000104738	MCM4	8:48872744-48890720	0.336343	0	5,00E-05	0.00305488	yes
ENSG00000213782	DDX47	12:12878850-12982915	0.484054	0	5,00E-05	0.00305488	yes
ENSG00000178660	ARMC10P1	3:94225609-94226464	0.289791	0	5,00E-05	0.00305488	yes
ENSG00000068878	PSME4	2:54091203-54307601	0.12927	0	0.00045	0.0201295	yes
ENSG00000063601	MTMR1	X:149861434-149933576	0.100801	0	5,00E-05	0.00305488	yes
ENSG00000155085	AKD1	6:109809108-110012420	0.239652	0	5,00E-05	0.00305488	yes
ENSG00000160948	VPS28	8:145648999-145653931	0.598512	0	5,00E-05	0.00305488	yes
ENSG00000138069	RAB1A	2:65283499-65357240	0.0738017	0	0.00015	0.00835	yes
ENSG00000058804	TMEM48	1:54231132-54304533	0.115543	0	5,00E-05	0.00305488	yes
ENSG00000131269	ABCB7	X:74273108-74376567	0.170756	0	5,00E-05	0.00305488	yes
ENSG00000185480	C12orf48	12:102513955-102591623	0.360215	0	0.00065	0.0266926	yes
ENSG00000145734	BDP1	5:70751441-70863649	0.0933384	0	5,00E-05	0.00305488	yes
ENSG00000115275	MOGS	2:74688183-74692537	0.117025	0	5,00E-05	0.00305488	yes
ENSG00000223658	AC011242.6.1	2:43864411-43995126	0.251771	0	0.0003	0.0147353	yes
ENSG00000156787	WDR67	8:124014799-124164393	0.160299	0	0.0001	0.00582558	yes
ENSG00000151503	NCAPD3	11:133938819-134117686	0.775431	0	0.00045	0.0201295	yes
ENSG00000196975	ANXA4	2:69686413-70053596	0.0578661	0	0.0002	0.0102245	yes

Supplementary Table 3. Erβ+ BC cell line Differential Promoter using at q-value ≤0.05.

gene_id	gene	locus	sqrt(JS)	test_stat	p_value	q_value	significant
ENSG00000234741	GAS5	1:173831289-173866494	0.136996	0	0.00105	0.0398523	yes
	HEATR1	1:236681299-236767804	0.121115	0	5,00E-05	0.00305488	yes
ENSG00000224831	RP11-651P23.4.1	3:149478891-149942977	0.418664	0	5,00E-05	0.00305488	yes
ENSG00000131263	RLIM	X:73805051-73834452	0.065689	0	0.00115	0.04175	yes
ENSG00000104738	MCM4	8:48872744-48890720	0.336343	0	5,00E-05	0.00305488	yes
ENSG00000213782	DDX47	12:12878850-12982915	0.484054	0	5,00E-05	0.00305488	yes
ENSG00000178660	ARMC10P1	3:94225609-94226464	0.289791	0	5,00E-05	0.00305488	yes
ENSG00000068878	PSME4	2:54091203-54307601	0.12927	0	0.00045	0.0201295	yes
ENSG00000063601	MTMR1	X:149861434-149933576	0.100801	0	5,00E-05	0.00305488	yes
ENSG00000155085	AKD1	6:109809108-110012420	0.239652	0	5,00E-05	0.00305488	yes
ENSG00000160948	VPS28	8:145648999-145653931	0.598512	0	5,00E-05	0.00305488	yes
ENSG00000138069	RAB1A	2:65283499-65357240	0.0738017	0	0.00015	0.00835	yes
ENSG00000058804	TMEM48	1:54231132-54304533	0.115543	0	5,00E-05	0.00305488	yes
ENSG00000131269	ABCB7	X:74273108-74376567	0.170756	0	5,00E-05	0.00305488	yes
ENSG00000185480	C12orf48	12:102513955-102591623	0.360215	0	0.00065	0.0266926	yes
ENSG00000145734	BDP1	5:70751441-70863649	0.0933384	0	5,00E-05	0.00305488	yes
ENSG00000115275	MOGS	2:74688183-74692537	0.117025	0	5,00E-05	0.00305488	yes
ENSG00000223658	AC011242.6.1	2:43864411-43995126	0.251771	0	0.0003	0.0147353	yes
ENSG00000156787	WDR67	8:124014799-124164393	0.160299	0	0.0001	0.00582558	yes
ENSG00000151503	NCAPD3	11:133938819-134117686	0.775431	0	0.00045	0.0201295	yes
ENSG00000196975	ANXA4	2:69686413-70053596	0.0578661	0	0.0002	0.0102245	yes

Supplementary material 1. (Erβ-) with selected differentially spliced genes (significant exon change) at an adjusted p-value≤0.05.

geneID	exonID	Dispersion	p-value	padjust	log2fold(E.MCF7TO/noE.MCF7TO)
ENSG00000005022	E007	0.0147055470136266	0	0	1.23698E+14
ENSG00000062716	E010	0.000760338066554652	0	0	0.188890253253861
ENSG00000062716	E015	0.000295163964057167	0	0	-0.104675732380832
ENSG00000062716	E020	0.000234579474480542	0	0	-0.0943458090571052
ENSG00000063177	E027	0.00891411550729496	0	0	0.0553905614482828
ENSG00000071082	E007	0.0144210199887183	0	0	0.0700737410615521
ENSG00000075624	E019	0.00455444228811321	0	0	-0.0644923967077791
ENSG00000109475	E002	0.00907538924849258	0	0	-0.0779595909508594
ENSG00000109475	E003	0.00546314773064575	0	0	-0.0656868304620877
ENSG00000111907	E024	0.00501112637065375	0	0	0.670055812031743
ENSG00000128609	E023	0.0311252561789971	0	0	1.65431E+14
ENSG00000128609	E024	0.0342686024037692	0	0	1.64574E+12
ENSG00000128641	E036	0.00603682819389217	0	0	-0.772440536766849
ENSG00000134333	E051	0.0447131631998469	0	0	0.438335394659404
ENSG00000140988+ENSG00000207405+ENSG00000255513+ENSG00000206811	E006	0.0348438876275743	0	0	0.0659382766423838
ENSG00000161016	E011	0.0106634397668682	0	0	-1.46983E+14
ENSG00000170889	E015	0.0109430196428278	0	0	0.112189002755497
ENSG00000206941+ENSG00000149273	E045	0.00157025629193141	0	0	-0.678273482281045
ENSG00000206941+ENSG00000149273	E048	0.000274692972329399	0	0	0.103829799147221
ENSG00000259001+ENSG00000252678	E001	0.000336625879494382	0	0	-0.0977170086797675
ENSG00000259001+ENSG00000252678	E002	0.000402833507196914	0	0	0.0422333243280596

Supplementary material 2. (Erβ+) with selected differentially spliced genes (significant exon change) at an adjusted p-value≤0.05.

geneID	exonID	Dispersion	pvalue	padjust	log2fold (with_E_5B12/no_E_5B12)
ENSG00000001631+ENSG00000243107	E029	0.0148526122373315	0	0	0.865040077151733
ENSG00000005700	E017	0.015898293844675	0	0	0.985529091793383
ENSG00000015153	E025	0.0150988457158889	0	0	0.953724663614144
ENSG00000031003	E022	0.00479163865635899	0	0	0.453751262738086
ENSG00000031003	E023	0.00336889926466394	0	0	0.554888688343613
ENSG00000033170	E016	0.0128992399400473	0	0	0.887002581926375
ENSG00000062716	E007	0.00171864314294444	0	0	0.386073943547944
ENSG00000062716	E008	0.00311704081967452	0	0	0.389584872758711
ENSG00000062716	E009	0.00190088447167943	0	0	0.289544969931232
ENSG00000062716	E015	0.00286459949724193	0	0	-0.30851002640611
ENSG00000062716	E016	0.00425652256325295	0	0	-0.305026234124033
ENSG00000062716	E017	0.00356481803370272	0	0	-0.335980847377032
ENSG00000062716	E018	0.00304181606709994	0	0	-0.3485959073919
ENSG00000062716	E019	0.00324253922385614	0	0	-0.362150147316055
ENSG00000062716	E020	0.00367726498344995	0	0	-0.366566409915974
ENSG00000063177	E018	0.0180659071070234	0	0	0.843553863217536
ENSG00000064419	E024	0.00509427483521333	0	0	0.501746375626447
ENSG00000064419	E025	0.00387433347133388	0	0	0.478221693303528

Supplementary material 2. Contd.

ENSG00000065559	E019	0.0047346883521754	0	0	-0.3892548049488
ENSG00000065833	E015	0.00638213123629173	0	0	0.643325750042247
ENSG00000065833	E016	0.00713957796110783	0	0	0.759016398261631
ENSG00000067225	E042	0.0203550550372131	0	0	110,983,190,688,681
ENSG00000068784	E013	0.00640862683473831	0	0	0.588913719428765
ENSG00000068784	E015	0.00490848909695605	0	0	0.530878094223965
ENSG00000068784	E016	0.0057852745979348	0	0	0.519287317592044
ENSG00000069020	E008	0.00799151004978879	0	0	0.621478519504317
ENSG00000070018	E037	0.00805879332343016	0	0	0.595272496170778
ENSG00000073921	E049	0.00239473416240031	0	0	0.309446443510305
ENSG00000073921	E051	0.00272295771365416	0	0	0.343864111345239
ENSG00000075415+ENSG00000212443	E022	0.00152116185896419	0	0	-0.268723862804229
ENSG00000075415+ENSG00000212443	E027	0.00179945036335906	0	0	-0.288335397378916
ENSG00000075415+ENSG00000212443	E028	0.00202095372077982	0	0	-0.296109552567584
ENSG00000075415+ENSG00000212443	E029	0.00143577303941475	0	0	-0.24120068415607
ENSG00000075415+ENSG00000212443	E030	0.00105444398579539	0	0	13,134,421,512,385
ENSG00000075415+ENSG00000212443	E031	0.00267605182220506	0	0	151,362,757,630,403
ENSG00000075415+ENSG00000212443	E033	0.00130871368914569	0	0	-0.267558868234547
ENSG00000075415+ENSG00000212443	E035	0.00143944275597132	0	0	-0.296074491191454
ENSG00000075415+ENSG00000212443	E036	0.00129603734922975	0	0	-0.298857446926918
ENSG00000077454+ENSG00000205307	E023	0.0129472931725166	0	0	0.937501657360712
ENSG00000080815	E025	0.00480298722660499	0	0	0.706879670381831
ENSG00000080815	E027	0.00688613769902666	0	0	0.79462813741011
ENSG00000080815	E028	0.00408933751263672	0	0	0.749617930831614
ENSG00000080815	E029	0.00382638192974004	0	0	0.633359607590383
ENSG00000082996	E047	0.00246856015033472	0	0	-0.302374927302371
ENSG00000083544	E010	0.00552035550206321	0	0	0.533214394515569
ENSG00000083544	E011	0.00432208822840506	0	0	0.444470863022972
ENSG00000084676	E008	0.0113129515504324	0	0	0.742318956342851
ENSG00000087206	E020	0.00610244376068148	0	0	0.511025033513407
ENSG00000088808	E052	0.00719296593685567	0	0	0.668990644176052
ENSG00000088808	E054	0.00875987294228689	0	0	0.741674100350788
ENSG00000089280	E039	0.00877485746148689	0	0	0.950136473057835
ENSG00000096746	E021	0.00812591195113053	0	0	117,421,325,765,102
ENSG00000099901	E031	0.00305195841705114	0	0	0.983367839219162
ENSG00000100941	E006	0.0206335783689133	0	0	11,693,174,711,559
ENSG00000101236	E002	0.0109607296293204	0	0	-0.531431758519199
ENSG00000101745	E026	0.00294489428487539	0	0	-0.334440688443602
ENSG00000104738	E015	0.0282472557665178	0	0	140,102,862,544,272
ENSG00000105176	E005	0.00561896839100511	0	0	0.52574623890558
ENSG00000105176	E006	0.00286607974584592	0	0	0.367031436796265
ENSG00000106462	E033	0.0068757494087806	0	0	0.549818906633497
ENSG00000107077+ENSG00000225489	E026	0.00784186024918819	0	0	0.672843604902932
ENSG00000107077+ENSG00000225489	E027	0.00862568037441722	0	0	0.742512595954742
ENSG00000109184	E048	0.00224309599285509	0	0	-0.241803208462778
ENSG00000109381	E032	0.00371487388040335	0	0	0.583175186270463
ENSG00000109670	E016	0.00288603265525797	0	0	0.300404990880352
ENSG00000110422+ENSG00000223134	E004	0.00598991292389415	0	0	0.528736156353527
ENSG00000110422+ENSG00000223134	E005	0.00287797953828343	0	0	0.412427216460058
ENSG00000111057	E018	0.00683382279778425	0	0	0.760593435804187
ENSG00000111371	E021	0.00172571436230456	0	0	0.285422353013813
ENSG00000111371	E022	0.00155471180260163	0	0	0.304166388693279

Supplementary material 2. Contd.

ENSG00000111371	E023	0.00121258732135146	0	0	0.382947940494121
ENSG00000111371	E024	0.00177287313329524	0	0	0.429832701792157
ENSG00000111371	E025	0.00673870834043841	0	0	0.702039552277118
ENSG00000111907	E024	0.00976249551033376	0	0	0.95357227979476
ENSG00000112851	E012	0.00232381998253929	0	0	0.490348157953446
ENSG00000112851	E013	0.00347586013562747	0	0	0.456944892834843
ENSG00000112893	E002	0.00317823266743197	0	0	0.492430327273209
ENSG00000112893	E003	0.00334748098552705	0	0	0.478273486420261
ENSG00000113643	E009	0.00225700164346493	0	0	0.396964193857818
ENSG00000113643	E010	0.00164000710748988	0	0	0.375127949708799
ENSG00000113643	E011	0.00187734282512215	0	0	0.370650749366847
ENSG00000113643	E013	0.00197520780529728	0	0	0.336827188191702
ENSG00000113643	E014	0.00227764236649451	0	0	0.320264470637763
ENSG00000113643	E029	0.00302403413694174	0	0	-0.351054375360191
ENSG00000113643	E032	0.00298329231540746	0	0	-0.350005775040493
ENSG00000113643	E033	0.00349229022771841	0	0	-0.391157974689573
ENSG00000114062	E002	0.00230106068441794	0	0	-0.327897303276237
ENSG00000114062	E016	0.00626506375865901	0	0	0.696718717460068
ENSG00000114062	E019	0.00727796516475796	0	0	0.670851499558417
ENSG00000114062	E020	0.00666757498149153	0	0	0.636686595223821
ENSG00000114062	E021	0.00294456300796946	0	0	0.619166465060826
ENSG00000114062	E022	0.0047453885810777	0	0	0.714667115837895
ENSG00000115053	E011	0.0174307505260267	0	0	0.956372961702347
ENSG00000115053	E013	0.0156844461759044	0	0	0.937540618360008
ENSG00000115109	E028	0.00236605544736564	0	0	-0.250438862134361
ENSG00000115109	E031	0.00250881042409631	0	0	0.817459630556919
ENSG00000115109	E033	0.0025187534100837	0	0	0.809713281704265
ENSG00000115109	E034	0.00252918259297746	0	0	0.767372756661953
ENSG00000115109	E035	0.00359136144768794	0	0	0.838983065087876
ENSG00000115109	E037	0.00379696035904649	0	0	0.750357993394999
ENSG00000115109	E038	0.00426666549720381	0	0	0.726117824877599
ENSG00000115109	E039	0.00494002930958599	0	0	0.703335604309351
ENSG00000115109	E040	0.00757486945342753	0	0	0.706793028856247
ENSG00000115109	E041	0.00560834337136765	0	0	0.653626614057737
ENSG00000115109	E044	0.00328057860142173	0	0	-0.431820103842305
ENSG00000115947	E027	0.00393841091047314	0	0	0.394557249798872
ENSG00000115947	E029	0.014359638168039	0	0	104,574,791,083,796
ENSG00000117868	E036	0.00332960442370569	0	0	0.409338273052346
ENSG00000120071	E013	0.00888880095647917	0	0	0.650660762856621
ENSG00000120071	E014	0.00359246151293338	0	0	-0.318044300187656
ENSG00000120438+ENSG00000206910+ENSG00000207392	E010	0.0143232369671226	0	0	155,002,194,713,202
ENSG00000120438+ENSG00000206910+ENSG00000207392	E011	0.0129360786438524	0	0	149,977,401,705,643
ENSG00000121741	E033	0.016180250688066	0	0	116,552,030,219,602
ENSG00000121989	E008	0.0136890307268968	0	0	0.836706580293666
ENSG00000122566	E004	0.00386124696591086	0	0	0.436883965780834
ENSG00000123066	E048	0.00161924012474238	0	0	0.757927671681865
ENSG00000128585	E014	0.00588291761095006	0	0	0.524555117399378
ENSG00000133316+ENSG00000222328	E054	0.00328592820223505	0	0	193,294,955,185,962
ENSG00000134108	E018	0.0249809117446936	0	0	146,428,235,596,882
ENSG00000134222	E014	0.0281004655331887	0	0	105,114,629,849,682

Supplementary material 2. Contd.

ENSG00000134419+ENSG00000170540+ENSG00000260342	E006	0.0015086881878047	0	0	0.210328225098956
ENSG00000134419+ENSG00000170540+ENSG00000260342	E009	0.00125801049076872	0	0	0.183361081044624
ENSG00000134419+ENSG00000170540+ENSG00000260342	E012	0.00102332113069753	0	0	0.176785804925073
ENSG00000134419+ENSG00000170540+ENSG00000260342	E013	0.00105111512651973	0	0	0.181709375341326
ENSG00000134419+ENSG00000170540+ENSG00000260342	E014	0.00100104210755371	0	0	0.191534482214219
ENSG00000134419+ENSG00000170540+ENSG00000260342	E024	0.00254824865578304	0	0	-0.226258218108307
ENSG00000134419+ENSG00000170540+ENSG00000260342	E025	0.00192833614241941	0	0	-0.274120390856121
ENSG00000134419+ENSG00000170540+ENSG00000260342	E026	0.0029083726592252	0	0	-0.272566284242751
ENSG00000134419+ENSG00000170540+ENSG00000260342	E028	0.00195603894854	0	0	-0.257212555473744
ENSG00000134419+ENSG00000170540+ENSG00000260342	E031	0.00176796633179388	0	0	-0.239324489671118
ENSG00000134758	E004	0.00675923797830655	0	0	0.610963881687083
ENSG00000134909	E040	0.0200404299607068	0	0	106,806,777,849,084
ENSG00000134909	E041	0.0164959595399716	0	0	0.984878220836215
ENSG00000135821	E023	0.00652021775202899	0	0	105,818,815,620,248
ENSG00000135829	E033	0.00821932343757605	0	0	0.737277622503705
ENSG00000136021	E007	0.00361595781726588	0	0	0.557131981141342
ENSG00000136021	E008	0.00533979098241841	0	0	0.73494290008176
ENSG00000136021	E009	0.00304139629458152	0	0	0.615794672529598
ENSG00000136021	E012	0.00386544580126646	0	0	0.466288891855297
ENSG00000136492	E007	0.00188835582227356	0	0	0.391173324979618
ENSG00000136492	E008	0.00200962436209889	0	0	0.399591328093891
ENSG00000136699	E029	0.0190478639124207	0	0	101,999,477,752,199
ENSG00000137776	E056	0.00459299355839347	0	0	0.453702984870248
ENSG00000138346	E033	0.00462626175999363	0	0	0.468281872765351
ENSG00000138376	E001	0.00262350485773804	0	0	-0.337749437085777
ENSG00000139597+ENSG00000139617+ENSG00000244754	E052	0.00349480013417698	0	0	0.432517682531562
ENSG00000139597+ENSG00000139617+ENSG00000244754	E053	0.00359886441124523	0	0	0.494331300388985
ENSG00000139597+ENSG00000139617+ENSG00000244754	E055	0.00324763120904549	0	0	0.485813293929168
ENSG00000139597+ENSG00000139617+ENSG00000244754	E060	0.00444709803935796	0	0	0.441739703725799
ENSG00000140396	E035	0.00393256248044073	0	0	0.59557174165757
ENSG00000140396	E036	0.00583870341377709	0	0	0.711897360662693
ENSG00000140396	E037	0.00713046629089922	0	0	0.698407826058823
ENSG00000140988+ENSG00000207405+ENSG00000255513+ENSG00000206811	E003	0.0048916412936643	0	0	-0.415414360922689
ENSG00000140988+ENSG00000207405+ENSG00000255513+ENSG00000206811	E004	0.00307670268826956	0	0	-0.307003565666234
ENSG00000140988+ENSG00000207405+ENSG00000255513+ENSG00000206811	E010	0.00450419661335741	0	0	0.833760532630357
ENSG00000140988+ENSG00000207405+ENSG00000255513+ENSG00000206811	E019	0.0158543700016008	0	0	0.753205094081311

Supplementary material 2. Contd.

ENSG00000140988+ENSG00000207405+ENSG00000255513+ENSG00000206811	E020	0.0049790530462157 2	0	0	100,042,608,404,356
ENSG00000143771	E019	0.00580520097517474	0	0	-0.5041393883146
ENSG00000143797	E026	0.00302999310622773	0	0	0.640097884989721
ENSG00000143797	E027	0.00296215379944861	0	0	0.801703371605897
ENSG00000143797	E028	0.00485746264540363	0	0	0.805104858767256
ENSG00000144036	E031	0.00546444790283555	0	0	0.502356874014504
ENSG00000144893	E006	0.00714940644321631	0	0	0.703862541135284
ENSG00000144935	E005	0.00600887095068696	0	0	0.581695853779746
ENSG00000144935	E006	0.00983878029905456	0	0	0.731898124492107
ENSG00000144935	E007	0.00452401038048158	0	0	0.520503916422583
ENSG00000145833+ENSG00000181904	E040	0.0168345677037895	0	0	103,200,845,784,413
ENSG00000146247	E041	0.00586044114182191	0	0	0.52106281646343
ENSG00000146247	E042	0.00621775722472215	0	0	0.527588929981903
ENSG00000146433	E006	0.00533184361147513	0	0	0.610389158762169
ENSG00000148334	E011	0.0170996959916619	0	0	0.98654097837898
ENSG00000148334	E012	0.00848502258958842	0	0	0.974722622401965
ENSG00000151292	E005	0.00238770654861252	0	0	0.447009837865454
ENSG00000151292	E006	0.00202357155119877	0	0	0.465030886133111
ENSG00000151292	E028	0.00513829761767353	0	0	-0.465483799698602
ENSG00000151466	E027	0.00671282815114215	0	0	0.547465798993094
ENSG00000153147	E023	0.002629307396414	0	0	0.341231102665827
ENSG00000153147	E024	0.00300773535154315	0	0	0.447890073547592
ENSG00000153147	E025	0.00256410264377357	0	0	0.376260255571111
ENSG00000155313	E004	0.00525180488605466	0	0	0.48555581709588
ENSG00000156011+ENSG00000244018	E005	0.00240412419426961	0	0	-0.281446493575435
ENSG00000156011+ENSG00000244018	E036	0.00689863041310596	0	0	0.597189559696127
ENSG00000156011+ENSG00000244018	E037	0.00547028551311603	0	0	0.617785477889397
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E016	0.00275340047434549	0	0	-0.38303906670846
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E030	0.00279960607623052	0	0	-0.408457196702123
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E031	0.00249095801198147	0	0	-0.431363680997929
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E032	0.00265598586358416	0	0	-0.370202825829498
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E033	0.00261195016889289	0	0	-0.414224393549785
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E034	0.0035776946911356	0	0	-0.422482033521442
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E037	0.00289849893607011	0	0	-0.406822286247417
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E038	0.00334877306929115	0	0	-0.41262176799185
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E044	0.00243606317247174	0	0	-0.467753878000779

Supplementary material 2. Contd.

ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E045	0.00345762767455533	0	0	-0.488182481794944
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E047	0.00472895529548749	0	0	0.435805022974449
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E049	0.00223173170649903	0	0	-0.498517262070707
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E052	0.00115189549505119	0	0	0.935724987229728
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E053	0.00137491633317313	0	0	0.691034785698287
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E054	0.0013975561364819	0	0	0.691073847995375
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E055	0.0014030438546344	0	0	0.68928044614001
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E056	0.00345856540979694	0	0	0.42783795705693
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E062	0.00245951902558292	0	0	-0.478019279500372
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E066	0.00327164795828963	0	0	-0.478799452946402
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E067	0.00350605632252888	0	0	-0.542164258795581
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E068	0.00515699694280688	0	0	-0.555088249985218
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E070	0.00203003877138065	0	0	-0.435397169463498
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E071	0.00380642449170613	0	0	-0.500010757652924
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E072	0.00191963090273959	0	0	-0.505857709281614
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E073	0.00202839488708531	0	0	-0.49491682781671
ENSG00000156976+ENSG00000200418+ENSG00000221420+ENSG00000238942+ENSG00000200320	E074	0.00179083331632279	0	0	-0.467558407351389
ENSG00000157107	E035	0.0064007240924872	0	0	0.741962665901774
ENSG00000157107	E036	0.00562177269318031	0	0	0.770913659764632
ENSG00000157107	E037	0.00798939232042488	0	0	0.653821831849602
ENSG00000161016	E027	0.00143609841806585	0	0	0.223354970922777
ENSG00000161016	E028	0.00279141700605149	0	0	0.338134525510153
ENSG00000161960+ENSG00000207152	E015	0.00721797104986602	0	0	0.865907954109202

Supplementary material 2. Contd.

ENSG00000163399	E022	0.0083465251426942	0	0	0.843606390042178
ENSG00000163527	E006	0.00184387440180405	0	0	0.241655481290931
ENSG00000163960	E001	0.00215200436341718	0	0	-0.22135756559334
ENSG00000163960	E011	0.00639705509140356	0	0	0.571678661886178
ENSG00000163960	E012	0.00700355643766015	0	0	0.728965166959754
ENSG00000163960	E013	0.00763245451948193	0	0	0.70157385489661
ENSG00000164898+ENSG00000146963	E009	0.00703578540729069	0	0	0.639251231077684
ENSG00000164898+ENSG00000146963	E011	0.00526343079222993	0	0	0.520757057855848
ENSG00000165322	E035	0.0016795661792196	0	0	0.461377389312872
ENSG00000165322	E036	0.00399604584864479	0	0	0.657302131724785
ENSG00000165322	E037	0.00802494090171092	0	0	0.653562893945478
ENSG00000165458	E030	0.0249500277094995	0	0	132,310,588,956,013
ENSG00000165458	E031	0.00997478438278549	0	0	127,435,235,988,798
ENSG00000166441+ENSG00000200983	E015	0.0127008087518793	0	0	116,897,921,519,154
ENSG00000166441+ENSG00000200983	E016	0.0127099666227081	0	0	10,988,627,912,785
ENSG00000166508	E026	0.0107712172226513	0	0	0.776312080658173
ENSG00000168234	E004	0.00799266268460179	0	0	0.613508174053404
ENSG00000168234	E017	0.00706536296863918	0	0	-0.54695122080361
ENSG00000168300	E015	0.0017959425060779	0	0	0.390997291771277
ENSG00000170571	E001	0.0039451242413448	0	0	-0.417074363235088
ENSG00000170571	E003	0.0024786415473555	0	0	-0.378387135132571
ENSG00000171132	E052	0.00739983119544396	0	0	-0.551582891048824
ENSG00000171456	E018	0.00777623331476681	0	0	102,567,100,973,926
ENSG00000174748	E033	0.00127123288715843	0	0	-0.175037895385076
ENSG00000174748	E034	0.00138393505120624	0	0	-0.27209252320527
ENSG00000174748	E035	0.0016730790310939	0	0	-0.252183364250906
ENSG00000175029	E029	0.0220718808482376	0	0	144,558,759,284,373
ENSG00000177600+ENSG00000199785	E014	0.00764616695379164	0	0	120,084,782,180,684
ENSG00000180573	E002	0.00221143715264896	0	0	0.2314128736999
ENSG00000180573	E004	0.00187775536125918	0	0	-0.213662493387306
ENSG00000182568+ENSG00000131374	E053	0.0154306133649054	0	0	0.916574865396555
ENSG00000184990	E016	0.0116844702678487	0	0	-0.715276347403957
ENSG00000185122	E011	0.00667644944274783	0	0	0.801609701353347
ENSG00000185122	E012	0.00869077668690626	0	0	0.75800522779588
ENSG00000188486	E002	0.002501804818815	0	0	-0.246722814514588
ENSG00000188486	E004	0.00465967186614012	0	0	0.660763644804424
ENSG00000188994	E014	0.00665864352520593	0	0	0.65137328793776
ENSG00000188994	E015	0.00664451917287129	0	0	0.609717313954432
ENSG00000188994	E028	0.00160039093305463	0	0	-0.174113784354817
ENSG00000196305	E018	0.0106171881395179	0	0	108,663,149,435,791
ENSG00000196323	E030	0.00264437984914019	0	0	0.32865159143385
ENSG00000196323	E033	0.00399908687553767	0	0	0.431836331628507
ENSG00000196507+ENSG00000172465	E007	0.00316649487083061	0	0	-0.334080891083039
ENSG00000196562	E047	0.00250212083003243	0	0	0.323648412117151
ENSG00000197323	E027	0.00145788738232438	0	0	0.333261343772383
ENSG00000197323	E028	0.00685203627436341	0	0	0.492107700640735
ENSG00000197323	E029	0.00139928204319553	0	0	0.450979355582536
ENSG00000197323	E030	0.00183679922227043	0	0	0.428880069858297
ENSG00000197409+ENSG00000196866	E001	0.00172784322418301	0	0	-0.225309873926191
ENSG00000197555	E008	0.0153539080947535	0	0	12,773,880,146,778
ENSG00000197555	E009	0.0138347962079179	0	0	118,283,138,392,894
ENSG00000197555	E010	0.00693835978409769	0	0	120,925,044,407,473

Supplementary material 2. Contd.

ENSG00000197555	E013	0.00795128120591918	0	0	123,557,179,344,223
ENSG00000198162	E004	0.00121154922061305	0	0	0.188235732881353
ENSG00000198162	E005	0.00153945011246118	0	0	0.215594596846282
ENSG00000198218	E011	0.00601535964135193	0	0	0.927244174929793
ENSG00000198218	E012	0.010144416450924	0	0	0.99351403204386
ENSG00000198363	E062	0.0043063981614386	0	0	0.553159372699458
ENSG00000198363	E063	0.00698653261112429	0	0	119,685,344,263,619
ENSG00000198363	E065	0.00349962656590338	0	0	0.661822814389495
ENSG00000198815	E019	0.00646598764600319	0	0	0.591750628117065
ENSG00000200259+ENSG00000201675+ENSG00000142541+ENSG00000202503+ENSG00000199631	E007	0.00779977227275279	0	0	0.718331946549828
ENSG00000200259+ENSG00000201675+ENSG00000142541+ENSG00000202503+ENSG00000199631	E008	0.00672936930243383	0	0	0.657555795340283
ENSG00000200463+ENSG00000179029	E002	0.0140137156483014	0	0	0.734536179627057
ENSG00000201129+ENSG00000143569	E048	0.0179917164916879	0	0	0.971192775925756
ENSG00000201808+ENSG00000242125+ENSG00000180198	E008	0.00226670779248858	0	0	22,836,062,220,647
ENSG00000201808+ENSG00000242125+ENSG00000180198	E011	0.00432483098196286	0	0	-405,813,828,106,499
ENSG00000201808+ENSG00000242125+ENSG00000180198	E048	0.00309453449892614	0	0	-465,753,535,559,348
ENSG00000204764	E054	0.00294568721692079	0	0	0.538583738222738
ENSG00000204764	E055	0.00348766153794291	0	0	0.593705920647754
ENSG00000204764	E056	0.00366240026945257	0	0	0.562404109509662
ENSG00000204842	E062	0.00368399705868222	0	0	0.515099379524403
ENSG00000204842	E064	0.00387052946002826	0	0	0.613313482164015
ENSG00000204842	E065	0.00804495883922123	0	0	0.669394526913633
ENSG00000207165+ENSG00000147403	E027	0.00169692053801728	0	0	0.617375222522592
ENSG00000212487+ENSG00000199437+ENSG00000233016	E004	0.00260564290326098	0	0	-0.59854798372185
ENSG00000212487+ENSG00000199437+ENSG00000233016	E005	0.00387684972471763	0	0	-0.607955590561775
ENSG00000212487+ENSG00000199437+ENSG00000233016	E007	0.0014830511297634	0	0	0.775476001818444
ENSG00000212487+ENSG00000199437+ENSG00000233016	E008	0.00622622892324662	0	0	0.55820527085017
ENSG00000212487+ENSG00000199437+ENSG00000233016	E009	0.00371124919009088	0	0	-0.594810946999488
ENSG00000215021+ENSG00000238795	E015	0.00507281302953734	0	0	0.415999483987769
ENSG00000215021+ENSG00000238795	E016	0.00462822138549259	0	0	0.446848247007522
ENSG00000215845+ENSG00000158769	E027	0.00236208580856115	0	0	0.272870922649241
ENSG00000234741+ENSG00000200729+ENSG00000200710+ENSG00000201692+ENSG00000202394+ENSG00000208313+ENSG00000200016+ENSG00000200954+ENSG00000206607	E048	0.00484913822193553	0	0	0.526348587933101
ENSG00000234741+ENSG00000200729+ENSG00000200710+ENSG00000201692+ENSG00000202394+ENSG00000208313+ENSG00000200016+ENSG00000200954+ENSG00000206607	E049	0.00365343042135109	0	0	0.524242044103493
ENSG00000241111+ENSG00000241572	E001	0.00389814531962009	0	0	0.353547187314918
ENSG00000246203+ENSG00000163374	E040	0.0061591245144764	0	0	0.441834318257926

Supplementary material 2. Contd.

ENSG00000251790+ENSG00000234912+ENSG00000129657	E012	0.00124939886827933	0	0	105,522,595,700,129
ENSG00000251790+ENSG00000234912+ENSG00000129657	E013	0.00461526473667066	0	0	0.48065205874616
ENSG00000251790+ENSG00000234912+ENSG00000129657	E039	0.012870208741594	0	0	-0.913531141877969
ENSG00000251790+ENSG00000234912+ENSG00000129657	E041	0.0169985343704731	0	0	-0.986462226319171
ENSG00000251790+ENSG00000234912+ENSG00000129657	E059	0.0111969522820811	0	0	-0.857889014514752
ENSG00000258508+ENSG00000198604	E047	0.00394845052589824	0	0	0.641303219622624
ENSG00000258941+ENSG00000150527+ENSG00000150526	E034	0.00786058374970822	0	0	0.829882498722205
ENSG00000258941+ENSG00000150527+ENSG00000150526	E035	0.0116095457657273	0	0	0.838478673247404
ENSG00000259001+ENSG00000252678	E001	0.00133201199818049	0	0	-0.132048289405817
ENSG00000259001+ENSG00000252678	E002	0.000823130909139826	0	0	0.0330027322644232
ENSG00000259932+ENSG00000259539+ENSG00000138606	E042	0.00779545365127278	0	0	0.513429082210731

academicJournals

Vol. 16(25), pp. 1428-1433, 21 June, 2017
DOI: 10.5897/AJB2017.16014
Article Number: 62E8DC264874
ISSN 1684-5315

http://www.academicjournals.org/AJB

African Journal of Biotechnology

Full Length Research Paper

Production potential and chemical composition of elephant grass (*Pennisetum purpureum* Schum.) at different ages for energy purposes

Ana Kesia Faria Vidal[1]*, Tatiane da Costa Barbé[2], Rogério Figueiredo Daher[1], Janeo Eustáquio Almeida Filho[1], Roberta Samara Nunes de Lima[3], Rafael Souza Freitas[1], Drieli Aparecida Rossi[4], Érik da Silva Oliveira[1], Bruna Rafaela da Silva Menezes[5], Geovana Cremonini Entringer[1], Wanessa Francesconi Stida Peixoto[1] and Sabrina Cassaro[1]

[1]Universidade Estadual do Norte Fluminense Darcy Ribeiro, CEP: 28013-602, Campos dos Goytacazes, RJ, Brazil.
[2]Faculdade de Ciências e Tecnologia de Unaí, CEP: 38610-000, Unaí, MG, Brazil.
[3]Universidade Federal de Alagoas, Campus Arapiraca, CEP: 57300-970, Arapiraca, AL, Brazil.
[4]Faculdade Venda Nova do Imigrante, CEP: 29375-000, Venda Nova do Imigrante, ES, Brazil.
[5]Universidade Federal Rural do Rio de Janeiro, CEP.: 23.897-000, Seropédica, RJ, Brazil.

Received 30 March, 2017; Accepted 18 May, 2017

The elephant grass has gained prominence as one of the main forage species used for biomass production. Therefore, the aim of this study was to identify elephant grass genotypes with high energy biomass production potential by evaluating morpho-agronomic and biomass quality. The following traits were evaluated in this study: dry matter yield (DMY), percentage of whole-plant dry matter (%DM), percentage of neutral detergent fiber (%NDF), percentage of acid detergent fiber (%ADF); percentage of cellulose (%CEL), percentage of lignin (%LIG), percentage of carbon (%C), percentage of nitrogen (%N), and carbon: nitrogen ratio (C: N). Five different production ages were evaluated, and significant differences were observed for the variable DMY. The harvests performed at 20 and 24 weeks of age, provided the best response for biomass energy production (DMY) from elephant grass, averaging 20.50 and 23.77 t.ha^{-1}. The genotypes that most stood out during the evaluation period at the five production ages were King Grass, Mole de Volta Grande, and Mercker 86 - México. Genotypes Mole de Volta Grande and King Grass are the most suitable for elephant grass breeding programs aimed at biomass energy production in the conditions of Campos dos Goytacazes - RJ, Brazil.

Key words: Bioenergy, biomass quality, carbon: nitrogen ratios, *Pennisetum purpureum* Shum, yield.

INTRODUCTION

The increasing search for alternatives to fossil fuels has become a critical issue for the future of the economic development of the planet. Amidst the possible solutions, biomass stands out as a medium- and long-term choice

*Corresponding author. E-mail: anakesia.vidal@hotmail.com.

(Goldemberg, 2009). Because biomass burning only recycles the CO_2 taken from the atmosphere by photosynthesis, it appears that, in the long term, this will be one of the safest energy alternatives, provided that it is produced efficiently (Kalt and Kranzl, 2011).

The elephant grass (*Pennisetum purpureum* Schum.) is triploid tropical forage of African origin, which presents high production capacity and quality dry matter accumulation. In this regard, elephant grass has gained prominence as one of the main forage species used in energy production (Morais et al., 2011).

According to Mazzarella (2006), the comparative advantages of elephant grass for biomass production in relation to other sources include: greater yield (around 45 t/DM/ha/year), shorter production time, better cash flow, possibility of total mechanization, renewable energy, and greater carbon assimilation. Since its products are not directly present in the human diet and because it is usable in its entirety, elephant grass is ahead of other grasses used as energy source.

Elephant grass species feature is a large genetic variability (Meinerz et al., 2011). For this reason, materials adapted to the different ecosystems of Brazil should be selected to broaden the understanding of the interrelationships among traits. As a result, the employment of elephant grass as a bioenergy source can be improved and this plant can be elevated to a prominent level in the sustainable diversification of the Brazilian energy matrix.

Therefore, aiming to investigate and understand the mechanisms to qualitatively and quantitatively increase the production of elephant grass for biomass energy generation, this study identified elephant grass genotypes with high biomass energy production by evaluating morpho-agronomic and biomass quality.

MATERIALS AND METHODS

This experiment was implemented on April 26, 2010, at the State Center for Research in Agro-Energy and Waste Utilization (PESAGRO-Rio), located in Campos dos Goytacazes, Northern Rio de Janeiro State, Brazil. The climate in the region is a tropical hot and humid Aw type, according to the Köppen (1948) classification, with dry winters and rainy summers.

The soil in the experimental area is classified as a Typic Dystrophic Yellow Latosol, according to the Brazilian soil taxonomy. Chemical analysis revealed the following soil composition: pH - 6.3; phosphorus - 5.0 mg/dm^3; potassium - 176 mg/dm^3; calcium - 2.6 cmol$_c$/dm^3; magnesium - 1.4 cmol$_c$/dm^3; aluminum - 0.1 cmol$_c$/dm^3; hydrogen + aluminum - 0.0 cmol$_c$/dm^3 and carbon - 1.26%. Total monthly precipitation rates recorded during the experimental period are presented in Table 1.

Six of the most productive genotypes of elephant grass, with the best biomass quality, were selected according to results obtained by Rossi (2010).

The experiment was undertaken in a randomized block statistical design with three replicates in an arrangement of plots subdivided into two factors: plots (genotypes) and subplots (number of harvests). Each plot consisted of 15-m row with 1-m spacing, and each subplot was composed of 3 m, in which only the two central meters of the row were considered for evaluation, disregarding half a meter from the borders of each row.

Table 1. Monthly precipitation recorded in the period from April 2010 to July 2011 in Campos dos Goytacazes - RJ.

Month	Precipitation (mm)	
	2010	2011
January	-	115.6
February	-	0.5
March	-	211.5
April	61.5	25.4
May	25.9	41.0
June	24.5	12.2
July	57.5	7.20
August	2.3	-
September	7.5	-
October	87.8	-
November	88.8	-
December	69.8	-
Total	424.88	413.40

Source: Meteorological Station of the State Center for Research in Agro-Energy and Waste Utilization (PESAGRO-Rio, Campos dos Goytacazes - RJ, Brazil).

After the crop establishment phase in the field, the plot was leveled-off on 08/05/2010, followed by topdressing with 30 kg.ha^{-1} ammonium sulfate and 20 kg.ha^{-1} potassium chloride. Seedlings were re-planted to recompose the plots, followed by topdressing on 10/31/2010. After this process and the reestablishment of the crop in the field, a second plot-leveling cut was made on December 10, 2010. After that date, harvests were performed in plots chosen at random 8, 12, 16, 20, and 24 weeks of age, on 02/02/2011, 03/02/2011, 04/05/2011, 04/26/2011, and 05/23/2011, respectively.

Before the harvest, the traits were assessed in samples of whole plants collected at random. Prior to being analyzed and evaluated in the laboratory, these samples were dried in a forced-air oven at 65°C for 72 h. After drying, samples were ground through a Wiley mill with 1-mm sieves and packed in glass bottles.

Plant dry matter yield in t.ha^{-1} (DMY), was estimated as the product between whole-plant fresh matter yield and the percentage of whole-plant dry matter; the obtained value was then converted to t.ha^{-1}.

The following biomass quality-related traits were evaluated: percentage of neutral detergent fiber (% NDF); percentage of acid detergent fiber (% ADF); percentage of cellulose (% CEL); and percentage of lignin (% LIG), were carried out at the Laboratory of Food Analysis of Embrapa Gado de Leite, in Juiz de Fora, MG, Brazil, by the method of near infrared reflectance (NIRS) (Van Soest, 1963), in a Perstorp analytical spectrometer, Silver Spring, MD, model 5000, coupled to a microcomputer equipped with ISI software version 4.1 (Infrasoft International, University, Park, PA). Preliminary evaluations were made for the calibration of the equipment using samples referring to the different cutting ages, and the final reading was performed using the wavelengths of 1100 to 2500 nanometers. Percentages of carbon (%C) and nitrogen (%N) and carbon:nitrogen (C:N) ratio were obtained using the CHNS/O Perkin Elmer (14.800) auto-analyzer.

Statistical analyses were performed using the GENES software (Computer Software for Genetics and Statistics) (Cruz, 2016), developed by the Federal University of Viçosa.

Table 2. Estimates of mean squares, means, and experimental coefficients of variation of six elephant grass genotypes in Campos dos Goytacazes, RJ, Brazil.

Age	Source of variation	G.L	[1]/DMY	MS					%C	%N	C:N
				%DM	%NDF	%ADF	%CEL	%LIG			
8 Weeks	Block	2	0.0022	3.2835	39.322	14.092	30.762	0.2875	0.3811	0.0603	3.1097
	Genotypes	5	1.7092**	2.9850^{ns}	3.7574*	4.7390*	2.6789*	0.7614^{ns}	1.1531^{ns}	0.5581**	14.9237**
	Resídue	10	0.2082	1.4912	0.9117	10.848	0.6401	0.301	1.2269	0.0356	1.1233
	Means	-	6.60	20.83	69.97	37.83	34.04	2.93	42.62	2.70	16.10
	CV(%)	-	6.90	5.86	1.48	2.52	2.35	18.66	2.59	6.97	6.58
12 Weeks	Block	2	0.8562	492.111	90.679	98.907	50.896	0.6425	0.7302	0.0028	1646.439
	Genotypes	5	97.2018**	267.1204**	2.5347 ^{ns}	4.9905 ^{ns}	2.0715 ^{ns}	3.2243 ^{ns}	0.8907 ^{ns}	0.2622**	46.970899*
	Resídue	10	15.875	1.926.121	35.579	48.686	64.736	12.611	14.048	0.0352	1354.692
	Means	-	15.55	44.36	79.71	48.27	40.36	7.22	43.24	1.72	25.8
	CV(%)	-	8.09	9.89	2.37	4.57	6.3	15.54	2.74	10.87	14.26
16 Weeks	Block	2	32.63	129.623	0.2686	16.38	0.0129	0.8109	51.914	0.0222	154.012
	Genotypes	5	46.5146**	29.1573 ^{ns}	4.0283**	7.3356*	3.5290*	1.77100*	1.1891 ^{ns}	0.0724*	21.7249*
	Resídue	10	20.313	95.384	0.6933	15.305	0.7229	0.3672	0.7119	0.0172	53.771
	Means	-	18.4	24.35	77.04	45.29	39.1	5.64	42.98	1.74	24.95
	CV(%)	-	7.74	12.68	1.08	2.73	2.17	10.74	1.96	7.53	9.29
20 Weeks	Block	2	10.707	32.944	10.088	0.435	0.1349	0.1783	0.5057	0.0151	4.7714
	Genotypes	5	38.6174**	16.2572**	19.5975**	17.0365**	9.2597**	2.0318*	4.5494*	0.0199*	2.8155 ^{ns}
	Resídue	10	17.432	24.816	0.6787	0.691	0.4255	0.3743	10.261	0.0058	1.7269
	Média	-	20.5	29.63	75.67	45.27	38.15	6.52	43.26	1.66	26.05
	CV(%)	-	6.46	5.31	1.08	1.83	1.7	1.7	2.34	4.60	5.04
24 Weeks	Block	2	257.048	150.056	13.526	20.377	0.8973	0.1293	18.227	0.0215	55.987
	Genotypes	5	46.0653**	14.0424 ^{ns}	5.9931 ^{ns}	6.9007 ^{ns}	5.2394 ^{ns}	0.3894 ^{ns}	6.8315**	0.0319 ^{ns}	6.5844 ^{ns}
	Resídue	10	73.237	105.926	31.398	35.031	17.127	0.5444	0.4975	0.0262	60.609
	Means	-	23.77	36.08	76.89	47.44	38.82	8.17	43.33	1.69	25.82
	CV(%)	-	11.38	9.01	2.3	3.94	3.37	9.02	1.62	9.5	9.53

1/DMY = plant dry matter yield, in t.ha−1; %DM = percentage of dry matter; %NDF = percentage of neutral detergent fiber; %ADF = percentage of acid detergent fiber; %CEL = percentage of cellulose; %LIG = percentage of lignin; %C = percentage of carbon; %N = percentage of nitrogen; and C:N = carbon:nitrogen ratio. ** = significant at the 1% probability by the F test * = significant at the 5% probability by the F test; ns = not significant. CV (%) = coefficient of variation.

RESULTS AND DISCUSSION

Variance analysis by production age

The variance analyses (P<0.01 and P<0.05) for the ages of the six genotypes evaluated (Table 2) revealed significant differences for almost all evaluated traits. This fact can be regarded as indicative of the phenotypic expression of the existing genetic variability, in the species and in the genotype collection evaluated (Pereira et al., 2008).

For the first production age with plants harvested at eight weeks, there were significant differences for the source of variation treatment (genotypes) and for most of the evaluated traits, except percentages of dry matter (%DM), %LIG, and %C. These findings confirm the existence of genetic variability between the treatments evaluated at this production age.

In the harvest performed at 12 weeks of age, there was an increase in relation to the previous harvest for the traits %DM, %LIG, and C: N, and also DMY, the trait of greatest importance for elephant grass, which averaged 6.60 and 15.55 t.ha^{-1} in the first and second harvests, respectively. Similar dry matter yields were obtained by Silva et al. (2014), who evaluated elephant grass hybrids with high production potential in Campos dos Goytacazes - RJ, Brazil.

These results can be explained by the climatic

conditions of the period with highest incidence of rainfall (20.51 to 212.68 mm) (Table 1), followed by the topdressing applied in the restoration of the plots, which might have provided a greater development of the genotypes.

This result proves the existing difference between harvests performed in different periods, indicating that plants may undergo changes in their structure and morphology due to adverse environmental conditions. According to Faria et al. (2009), the genotype × environment interaction is of great importance in plant breeding; however, it compromises the identification of genotypes superior for different environments. At the harvest performed at 16 weeks of age, there were significant differences between genotypes for the majority of the evaluated traits, except %DM and %C.

For the fourth production age, at 20 weeks, there were significant differences between genotypes for most of the evaluated traits, except C: N ratio. The C: N ratio values found here were lower than the 80 to 156 obtained by Morais et al. (2009); the highest C: N ratio found in the present study was 26.05. This low ratio was due to the favorable climatic conditions for elephant grass growing (Flores et al., 2013).

At the fifth (and last) harvest performed at 24 weeks of age, the %DM, %NDF, %ADF, %CEL, %LIG, %N, and C:N traits showed a non-significant effect, demonstrating that the constitution of the walls of the tissues are similar across the genotypes.

In summary, at all five production ages, significant differences were observed for the variable DMY. Evaluating harvests individually by variance analysis, we observe that the fourth and fifth production ages (20 and 24 weeks of age) yielded the best responses for elephant grass biomass production (DMY), which averaged 20.50 and 23.77 $t.ha^{-1}$.

These results are similar to those found by Morais et al. (2009), who evaluated the production potential of different elephant-grass genotypes at different production ages (9, 18, and 24 weeks) and observed an average DMY of 29.5 $t.ha^{-1}$. Our results are also similar to the 11.15 to 23.08 $t.ha^{-1}$ found by Rocha et al. (2015), who evaluated 73 elephant grass genotypes in harvests performed at 24 weeks.

This demonstrates that elephant grass can be harvested more than two times per year, as typical, for biomass production as long as ideal conditions of planting and preservation of the crop in the field are provided.

Comparison of genotype means by production age (Tukey's test)

Based on Tukey's mean comparison test at the 5% probability level (Table 3), the estimated means for the traits evaluated in six genotypes at five different production ages showed that, at the first production age for the trait DMY, the best response was observed for genotype Pusa Napier no. 1, whose average dry matter yield was 7.27 $t.ha^{-1}$. For NDF, ADF, CEL, and %LIG traits, genotype Mole de Volta Grande had a better response than the others. At 12 weeks of age, King Grass was the genotype that most stood out for average DMY, %NDF, %ADF, %CEL, and %LIG, revealing 26.54 $t.ha^{-1}$, 80.03%, 50.44%, 40.95%, and 8.80%, respectively. For the C: N trait, genotype Mercker 86 - México showed the best response: 31.25.

As in the first harvest, in the harvest performed at 16 weeks of age, genotype Pusa Napier no. 1 had the best average production, 24.18 $t.ha^{-1}$. For the other traits of high importance for energy production, such as NDF and %CEL, genotype Mole de Volta Grande obtained 78.65% NDF and 40.56% CEL, respectively. As for the %ADF and %LIG traits, genotype Mercker 86-México averaged 47.16 and 6.22%, respectively. At the harvest performed at 20 weeks of age, genotype Mole de Volta Grande showed the best response, with a DMY of 23.62 $t.ha^{-1}$, 79.29% NDF, and 48.83% ADF. Its cellulose percentage was 40.63%, corroborating Kannika et al. (2011), who reported an increase in %CEL as plants grew older.

At the fifth and last production harvest, which was performed at 24 weeks of age, the genotype that most stood out for DMY and %LIG was Mercker 86 - México, which averaged 28.87 $t.ha^{-1}$ and 8.62% for the respective variables. For traits like NDF, ADF and %CEL, all genotypes except Pusa Napier no. 1 displayed good results. For C: N, as in the first harvest, genotype Cubano de Pinda showed a higher mean than the others, averaging 28.25.

Based on Tukey's mean comparison test at the 5% probability level, the materials that most stood out at the five different production ages were genotypes King Grass, Mole de Volta Grande, and Mercker 86 - México. These results are much higher as compared with those found by Rossi (2010) for these same materials, considering that they were obtained at the production age of 10 months; higher than those reported by Morais et al. (2009).

Combined analysis of variance

Table 4 depicts the mean squares for the effects of plot (genotypes), error A, associated with the plot effect; subplot (harvest), error B, associated with the subplot effect; interaction between plot and subplot; and error C, associated with the interaction effect, for the nine traits assessed in this study.

Most of the traits showed significant differences for the source of variation Genotype, except %DM, %C, and C: N ratio, indicating a difference between the genotypes. This result proves the distinction between the harvests performed in different periods, revealing that plants may

Table 3. Mean values for morpho-agronomic and biomass quality traits of six elephant-grass genotypes. Campos dos Goytacazes, RJ, Brazil.

Age (Weeks)	Genotypes	1/DMY	Mean values					%C	%N	C:N
			%DM	%NDF	%ADF	%CEL	%LIG			
8	Cubano de Pinda	7.23^{a}	21.34^{a}	69.98^{ab}	37.22^{ab}	33.70^{ab}	2.59^{a}	41.65^{a}	2.31^{b}	18.00^{a}
	Mercker 86 – México	6.25^{ab}	19.50^{a}	69.06^{ab}	37.23^{ab}	33.82^{ab}	2.48^{a}	42.32^{a}	2.48^{b}	17.11^{ab}
	Pusa Napier nº1	7.27^{a}	19.90^{a}	68.03^{b}	36.38^{b}	32.52^{b}	2.93^{a}	42.86^{a}	2.45^{b}	17.57^{a}
	Mole de Volta Grande	6.81^{a}	21.70^{a}	71.37^{a}	39.23^{a}	35.32^{a}	3.63^{a}	43.25^{a}	2.50^{b}	17.28^{a}
	P - 241 – Piracicaba	5.26^{b}	21.95^{a}	71.09^{a}	37.86^{ab}	34.37^{ab}	2.51^{a}	43.25^{a}	3.06^{b}	14.11^{bc}
	King Grass	6.80^{a}	20.59^{a}	70.28^{ab}	39.04^{ab}	34.53^{ab}	3.47^{a}	42.41^{a}	3.41^{a}	12.53^{c}
	Cubano de Pinda	12.08^{c}	44.38^{b}	79.90^{a}	47.59^{a}	40.49^{a}	6.47^{a}	43.84^{a}	2.29^{a}	19.19^{b}
	Mercker 86 – México	12.97^{c}	38.60^{bc}	77.60^{a}	47.13^{a}	38.84^{a}	6.53^{a}	43.80^{a}	1.44^{b}	31.25^{a}
12	Pusa Napier nº1	16.69^{b}	43.65^{bc}	79.58^{a}	48.78^{a}	41.00^{a}	8.21^{a}	42.37^{a}	1.62^{b}	26.30^{ab}
	Mole de Volta Grande	11.31^{c}	31.35^{c}	79.80^{a}	48.61^{a}	40.86^{a}	6.96^{a}	43.02a	1.72^{b}	25.12^{ab}
	P - 241 – Piracicaba	13.75^{bc}	48.96^{ab}	78.95^{a}	47.06^{a}	40.01^{a}	6.34^{a}	43.14^{a}	1.71^{b}	25.19^{ab}
	King Grass	26.54^{a}	59.23^{a}	80.03^{a}	50.44^{a}	40.95^{a}	8.80^{a}	43.27^{a}	1.563^{b}	27.75^{ab}
	Cubano de Pinda	15.29^{c}	23.30^{a}	77.07^{ab}	44.55^{ab}	38.70^{ab}	5.33^{ab}	42.90^{a}	1.76^{ab}	24.38^{ab}
	Mercker 86 – México	15.37^{c}	29.31^{a}	78.07^{ab}	47.16^{a}	40.15^{ab}	6.22^{a}	43.50^{a}	1.72^{ab}	25.24^{ab}
16	Pusa Napier nº1	24.18^{a}	22.16^{a}	75.90^{b}	44.11^{ab}	37.99^{b}	5.93^{a}	41.76^{a}	1.82^{a}	23.01^{b}
	Mole de Volta Grande	19.97^{b}	26.79^{a}	78.65^{a}	46.82^{a}	40.56^{a}	6.03^{a}	43.44^{a}	1.84^{a}	23.62^{ab}
	P - 241 – Piracicaba	14.49^{c}	20.96^{a}	75.72^{b}	43.26^{b}	37.98^{b}	4.21^{b}	42.98^{a}	1.84^{a}	23.28^{b}
	King Grass	21.13^{ab}	23.56^{a}	76.86^{ab}	45.86^{ab}	39.21^{ab}	6.12^{a}	42.72^{a}	1.44^{b}	30.19^{a}
	Cubano de Pinda	14.82^{c}	27.92^{b}	76.45^{b}	45.77^{b}	38.67^{bc}	6.50^{ab}	43.48^{ab}	1.80^{a}	24.31^{a}
	Mercker 86 – México	20.66^{ab}	31.02^{ab}	76.67^{b}	46.64^{ab}	39.02^{ab}	7.23^{ab}	45.09^{a}	1.68^{b}	26.74^{a}
20	Pusa Napier nº1	23.05^{a}	27.85^{b}	71.64^{c}	42.13^{d}	35.45^{d}	5.80^{b}	42.79^{ab}	1.58^{b}	27.14^{a}
	Mole de Volta Grande	23.62^{a}	33.64^{a}	79.29^{a}	48.83^{a}	40.63^{a}	7.75^{a}	44.05^{ab}	1.70^{ab}	25.94^{a}
	P - 241 – Piracicaba	17.63^{bc}	27.91^{b}	74.39^{b}	43.39^{cd}	37.17^{cd}	5.65^{b}	42.58^{ab}	1.63^{ab}	26.07^{a}
	King Grass	23.25^{a}	29.43^{ab}	75.58^{b}	44.86^{bc}	37.95^{bc}	6.22^{ab}	41.56^{b}	1.59^{ab}	26.12^{a}
	Cubano de Pinda	22.21^{abc}	34.30^{a}	76.64^{a}	46.87^{a}	38.26^{a}	7.94^{a}	43.97^{ab}	1.55^{a}	28.25^{a}
	Mercker 86 – México	28.87^{a}	34.57^{a}	77.55^{a}	48.97^{a}	39.84^{a}	8.62^{a}	42.71^{bc}	1.64^{a}	26.29^{a}
24	Pusa Napier nº1	20.37^{bc}	34.67^{a}	75.52^{a}	46.07^{a}	37.24^{a}	8.22^{a}	43.9^{ab}	1.73^{a}	25.67^{a}
	Mole de Volta Grande	19.11^{c}	39.96^{a}	79.41^{a}	49.72a	40.92^{a}	8.53^{a}	42.05^{bc}	1.77^{a}	23.73^{a}
	P - 241 – Piracicaba	24.51^{abc}	37.02^{a}	75.98^{a}	46.46^{a}	38.44^{a}	7.69^{a}	41.62^{c}	1.61^{a}	25.83^{a}
	King Grass	27.58^{ab}	35.97^{a}	76.24^{a}	46.56^{a}	38.25^{a}	8.01^{a}	45.71^{a}	1.82^{a}	25.15^{a}

1/DMY = plant dry matter yield, in t.ha−1; %DM = percentage of dry matter; %NDF = percentage of neutral detergent fiber; %ADF = percentage of acid detergent fiber; %CEL = percentage of cellulose; %LIG = percentage of lignin; %C = percentage of carbon; %N = percentage of nitrogen; and C:N = carbon:nitrogen ratio. Means followed by common letters do not differ statistically by Tukey's test at the 5% probability level.

undergo changes in their structure and morphology stemming from adverse environmental conditions. The source of variation harvests was also significant for most traits, except %C, indicating the presence of variability, which is paramount to prove the distinction between harvests performed during this study.

The only trait that did not show significance for the genotype × harvest interaction was %CEL. The significant interaction between genotypes and harvests indicates that there was a change in the classification of genotypes in different harvests. Sousa et al. (2016) evaluated elephant grass genotypes and also observed that, some were influenced by the evaluation harvest which is related to environmental conditions (temperature, luminosity, rainfall distribution), prevailing during the crop growth period.

The significant interaction between genotypes and harvests indicates that, the genotypes' response is not the same over successive harvest that is, there are differences between the genotype means, or in the

Table 4. Summary of the combined analysis of variance of the morpho-agronomic and biomass quality traits of six elephant-grass genotypes. Campos dos Goytacazes, RJ, Brazil.

Source of variation	D.F.	1/DMY	QM					%C	%N	C:N
			%DM	%NDF	%ADF	%CEL	%LIG			
Block	2	10.2986	109.924	37.186	96.758	37.334	0.2224	14.073	0.0158	12.246
Genotype (G)	5	88.7298**	31.1407 ns	20.3212**	22.1672**	13.0456*	3.6635**	1.297 ns	0.0777*	14.4680 ns
Error A	10	19.928	93.764	25.71	29.999	26.422	0.5324	13.164	0.0205	59.281
Cortes (C)	4	766.5816**	1592.5911**	220.9253**	306.4421**	103.9409**	71.7133**	1.643 ns	3.6147**	331.9834**
Error B	8	51.496	71.186	23.471	20.644	13.693	0.4565	18.060	0.0265	73.257
Genotype × Cut	20	35.3446**	74.6054**	4.1428**	4.4634*	2.4333 ns	1.1286*	3.3291**	0.2167**	19.6378**
Error C	40	2.725.319	84.971	16.458	21.263	18.331	0.5789	0.8877	0.0249	54.768

DF = degree of freedom; 1/DMY= plant dry matter yield, in t.ha−1; %DM = percentage of dry matter; %NDF = percentage of neutral detergent fiber; %ADF = percentage of acid detergent fiber; %CEL = percentage of cellulose; %LIG = percentage of lignin; %C = percentage of carbon; %N = percentage of nitrogen; and C:N = carbon:nitrogen ratio. ** = significant at the 1%probability by the F test * = significant at the 5% probability by the F test; ns = not significant.

classification of their responses, along the five harvests.

Conclusion

For the five production ages, significant differences were detected for variable DMY. The fourth and fifth ages provided the best responses.

Genotypes King Grass, Mole de Volta Grande, and Mercker 86 - México showed to be the most promising, and thus can ensure the use of elephant grass as an alternative energy source, with low energy production costs for the region of Campos dos Goytacazes.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

REFERENCES

Cruz CD, (2016). Genes Software – extended and integrated with the R, Matlab and Selegen. Acta Sci. Agron. 38(4):547-552.

Faria AP, Cirino A. P, Buratto JS, Silva CFB, Destro D (2009). Interação genótipo x ambiente na produtividade de grãos de linhagens e cultivares de feijão. Acta Sci. Agron 31(4):579-585.

Flores RAS, Urquiaga BJR, Alves LS, Collier J B, Zanetti R M, (2013). Nitrogênio e idade de corte na qualidade da biomassa de capim-elefante para fins agroenergéticos cultivado em Latossolo. Semina: Ciências Agrarias 34(1):127-136.

Goldemberg J (2009). Biomassa e energia. Quím. Nova 32(3):582-587.

Kalt G, Kranzl L (2011). Assessing the economic efficiency of bioenergy technologies in climate mitigation and fossil fuel replacement in Austria using a techno-economic approach. Appl. Energy 88:3665-3684.

Kannika R, Yasuyuki L, Kunn K, Pichit P, Prapa S, Vittaya P, Pilanee V, Ganda N, Sayan T (2011). Effects of inter-cutting interval on biomass yield, growth components and chemical composition of napiergrass (*Pennisetum purpureum* Schum.) cultivars as bioenergy crops in Thailand. Grassl. Sci. 57:135-141.

Köppen W (1948). Climatologia: con un estudio de los climas de la tierra. México D.F, Fondo de Cultura Econômica. 479 p.

Mazzarella VNG, (2006). Capim elefante como fonte de energia no Brasil: realidade atual e expectativas. Workshop Madeira Energética. Rio de Janeiro, 2006.

Meinerz GR, Olivo CJC, Agnolin A, Dullius AP, Moraes RS, Mombach G, Foletto V, Machado PR (2011. Produção e valor nutritivo da forragem de capim-elefante em dois sistemas de produção. R. Bras. Zootec. 40(12):2673-2680.

Morais RF, Quesada DM, Reis VM, (2011). Contribution of biological nitrogen fixation to Elephant grass (*Pennisetum purpureum* Schum.) Plant Soil 349:1-12.

Morais RF, Souza BJ, Leite JM, Soares LHB, Alves BJR, Boddey RM, Urquiaga S (2009). Elephant Grass genotypes for bioenergy production by direct biomass combustion. Pesqui. Agropecu. Bras. 44(2):133-140.

Pereira AV, Machado MA, Azevedo ALS, Nascimento S, Campos AL, Lédo FJS (2008). Diversidade genética entre acessos de capim-elefante obtida com marcadores moleculares. R. Bras. Zootec. 37:1216-1221.

Rocha AS, Daher RF, Gravina GA, Pereira AV, Rodrigues EV, Viana AP, Silva VQR, Junior ATA, Novo AAC, Oliveira MLF, Oliveira ES (2015). Comparison of stability methods in elephant-grass genotypes for energy purposes. Afr. J. Agric. Res. 10(47):4283-4294.

Rossi DA (2010) Avaliação morfoagronômica e da qualidade da biomassa de acessos de capim-elefante (*Pennisetum purpureum* Schum.) para fins energéticos no Norte Fluminense. Dissertation (Master in Plant Production) - Campos dos Goytacazes -RJ, State University of Norte Fluminense Darcy Ribeiro – UENF, 66p.

Silva VQR, Daher RF, Gravina GA, Lédo FJS, Tardin FD, Souza MC (2014). Capacidade combinatória de capim elefante com base em caracteres morfoagronômicos. B. Indústr. Anim. 71(1):63-70.

Sousa LB, Daher RF, Menezes BRS, Rodrigues EV, Tardin FD, Gravina AG, Pereira AV (2016). Qualidade da biomassa em híbridos de capim-elefante para fins energéticos. Agrária - Revista Brasileira de Ciências Agrárias. 11(2):85-91.

Van Soes PJ, (1963). Use of detergents in the analysis of fibrous feeds. A rapid method for the determination of fiber and lignin. J. Assoc. Off. Agric. Chem. 46(5):829-835.

African Journal of Biotechnology

Related Journals Published by Academic Journals

- *Biotechnology and Molecular Biology Reviews*
- *African Journal of Microbiology Research*
- *African Journal of Biochemistry Research*
- *African Journal of Environmental Science and Technology*
- *African Journal of Food Science*
- *African Journal of Plant Science*
- *Journal of Bioinformatics and Sequence Analysis*
- *International Journal of Biodiversity and Conservation*

academicJournals