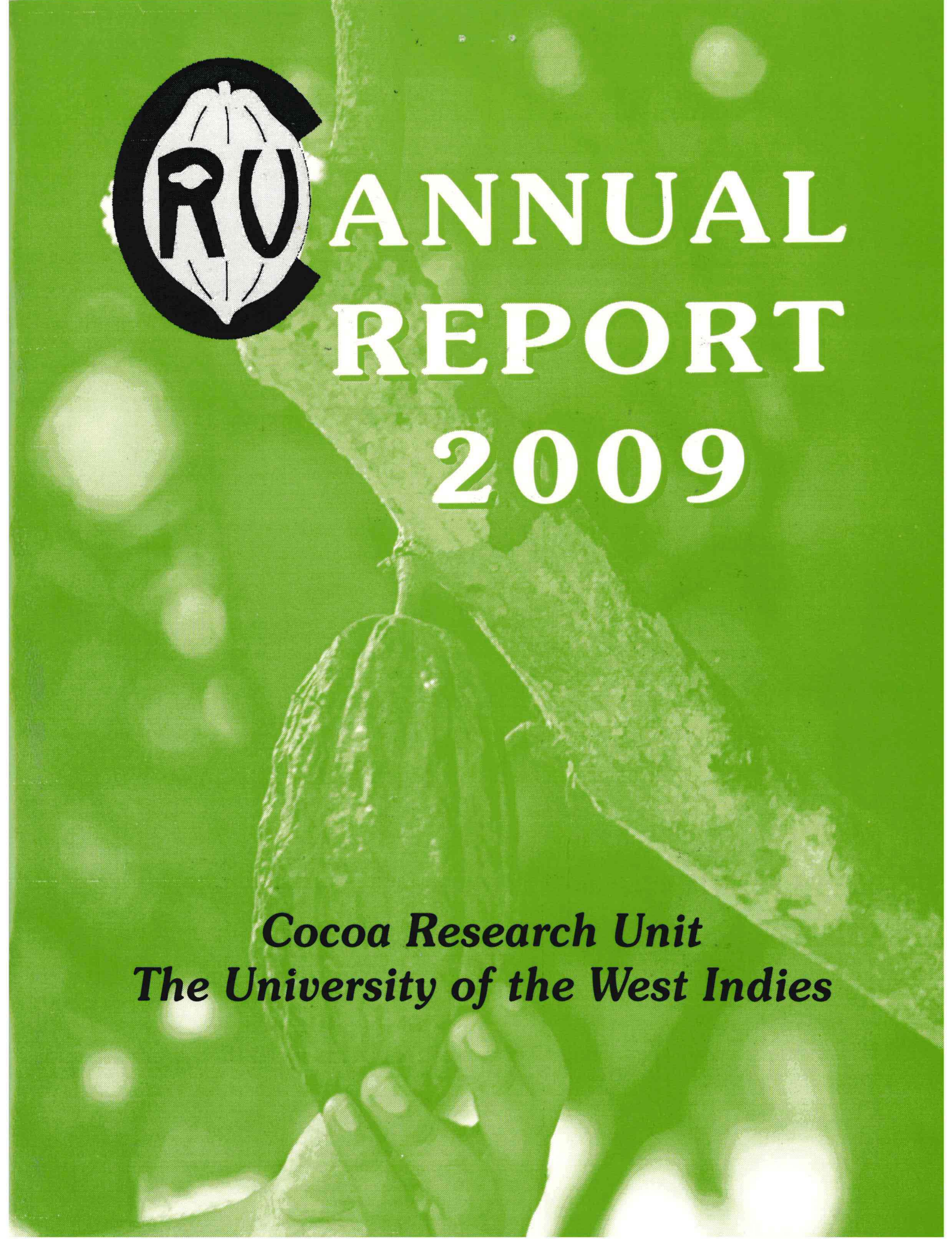

ANNUAL REPORT 2009

Cocoa Research Unit
The University of the West Indies

Annual Report 2009. St. Augustine, Trinidad and Tobago: Cocoa Research Unit, the University of the West Indies. 85 pp.

The work of CRU is made possible by support from

CRA Ltd. research towards sustainable cocoa — Cocoa Research Association, UK

CIRAD — Centre de coopération internationale en recherche agronomique pour le développement, France

Ministerie van Landouw, Natuur en Voedselkwaliteit, Holland

United Nations Common Fund for Commodities

Lindt MAÎTRE CHOCOLATIER — Lindt & Sprüngli (International) AG, Switzerland

Guittard Chocolate Company, Burlingame, USA

Cadbury — Cadbury Ltd., UK

Masterfoods — Masterfoods, UK

Bioversity International — Bioversity International

VALRHONA — Valrhona, France

Ministry of Agriculture, Land and Marine Resources, Government of the Republic of Trinidad and Tobago

USDA — United States Department of Agriculture, USA

WCF — World Cocoa Foundation, USA

The University of the West Indies, Trinidad and Tobago

University of Reading — The University of Reading, UK

International Cocoa Germplasm Database

UHH Universität Hamburg — The University of Hamburg, Germany

Queensland Government — Queensland Department of Primary Industries & Forestry, Australia

Towson University, Maryland, USA

Cocoa Research Unit
The University of the West Indies
St. Augustine, Trinidad and Tobago

Tel: +1 868 662 8788
+1 868 662 2002 Ext. 2115
Fax:+1 868 662 8788
E-mail: cru@sta.uw.edu

Cover photograph. Trinitario cocoa pod at the ICG,T being examined by Valmiki Singh

Annual Report 2009

Cocoa Research Unit
The University of the West Indies
St. Augustine, Trinidad and Tobago
2010

DNA fingerprinting of *Theobroma cacao* accessions with SSR primers

A.A. Sankar, M. Boccara, and L.A. Motilal

Introduction

Participation by CRU in the USDA Fingerprinting Project, with the goal to fingerprint *Theobroma cacao* accessions held at genebanks located in the Americas, was initiated in 2001. It was a project of enormous scope involving the use of 15 simple sequence repeat (SSR) primers to establish a genotypic fingerprint for each of the accessions to be analysed. In accordance with the project mandate, a representative sample of DNA from the most original tree of each accession in the ICG,T was sent to Beltsville, Maryland (USA) for analysis. The chosen mode of SSR analysis using a capillary electrophoresis system is well-designed for high-throughput processing; however the results require a considerable investment of time to be transformed into usable data. It is not remarkable, therefore, given the constraints of time and budget that not all of the approximately 2,300 DNA samples sent from CRU, Trinidad to Beltsville have yet been completed. In 2008, several hundred samples remained to be fingerprinted. Some of them represented accessions that had not yet been processed, while others were samples that proved challenging to analyse, possibly due to variable DNA concentration or poor sample quality. The materials and resources available at the Beltsville Agricultural Research Center (BARC) Sustainable Perennial Crops Lab (SPCL) headed by Dr. Lyndel Meinhardt make it an ideal location for processing the outstanding DNA samples at minimal cost to generate the important data required to continue characterisation, verification and genetic diversity studies. At the invitation of SPCL geneticist Dr. Zhang, with the approval of Dr. Meinhardt, Antoinette Sankar was therefore assigned to spend six months in Beltsville, to learn to operate the Beckman Coulter CEQ 8000 Genetic Analysis System, to perform Polymerase Chain Reaction (PCR) and to conduct analyses of as many samples as possible. Upon arrival in Beltsville, the list contained 866 samples to be analysed (D. Zhang pers comm). Some accessions on the list were classified as "missing" which was an indication that they could have been misplaced during laboratory relocation.

The CEQ Genetic Analysis system was used to perform capillary electrophoresis and analysis of PCR fragments produced by the amplification of the cacao DNA. PCR amplification was done using the SSR primers officially established for the USDA Fingerprinting Project (Saunders et al., 2004). The resultant fingerprinting profiles were exported as numeric allele data. Approximately 400 accessions were processed and data obtained for a subset of that number will be used for verification and genetic diversity studies. The work program consisted of training to use the CEQ system, cataloguing of samples, trial runs to optimise the DNA dilution factor, followed by systematic analysis of the cacao DNA samples alongside known cacao DNA controls.

Materials and Methods

All available DNA samples were sorted to separate those that had been completed from the ones which needed to be analysed. All samples that were not included on the outstanding list of 866 were classified as completed and separated from the rest. Trial runs were done to determine the

DNA dilution required for optimal results after which serial dilutions (up to the working concentration of 1:1000) of the stock DNA were prepared in 96-well PCR sample plates. From these plates, aliquots of DNA were removed by pipetting into new PCR plates for PCR amplification, after the samples had been vortexed and spun down to mix and collect the DNA at the bottom of the well. A sub-optimal number of positive DNA controls were used in the PCR due to short supply of this material.

Table 1. Fluorescent-labelled SSR primer pairs.

SPCL title*	CIRAD nomenclature	Expected Range
Primer 9	**mTcCIR1**	122-156
Primer 3	**mTcCIR6**	218-246
Primer 1	mTcCIR7	150-167
Primer 7	**mTcCIR8**	276-322
Primer 10	mTcCIR11	286-322
Primer 11	**mTcCIR12**	187-273
Primer 12	**mTcCIR15**	221-256
Primer 16	mTcCIR18	329-354
Primer 4	**mTcCIR22**	271-288
Primer 5	mTcCIR24	184-201
Primer 21	**mTcCIR26**	267-313
Primer 15	**mTcCIR33**	270-344
Primer 14	**mTcCIR37**	132-184
Primer 22	mTcCIR40	261-288
Primer 25	**mTcCIR60**	187-210

*Primer pairs numbered according to the system used at SPCL

PCR with Fluorescent-labelled SSR primers (Table 1) was done in accordance with SPCL protocol (S. Pinney, pers comm). The PCR master mix was prepared on ice using the composition shown in Table 2. All of the components, except the polymerase (Amplitaq Gold®) were thawed and vortexed briefly to collect the tube contents before pipetting. The reaction mix was pulsed briefly before addition of the polymerase. Both before and after adding primers to the PCR mix, they were protected from light to prevent primer degradation. The polymerase was the last component to be added and was spun down in an Eppendorf Centrifuge 5417R (refrigerated) for one minute then added to the reaction mix. The mix was incubated on ice while being added to the DNA template.

For PCR, 96-well plates of DNA were amplified according to the program shown in Table 3 (or a slight variant of it) using one of three thermal cyclers; these were two MJ Research PTC-200 Peltier Thermal Cyclers (one a Gradient cycler) and one AB (Applied Biosystems) GeneAmp PCR System 9700 cycler. An annealing temperature of 51°C was used for all primers as recommended.

PCR products were stored at -5 °C until they were prepared for electrophoresis in the Beckman Coulter CEQ 8000 Genetic Analysis System (CEQ 8000 System) to separate PCR fragments. To prepare samples for loading into the instrument, a Sample Loading Solution (SLS)

Table 2. PCR mix composition.

Component	Volume (µl)
Water	0.02
10X Amplitaq Gold® Buffer	1.0
10 mg/ml BSA	1.0
25mM $MgCl_2$	1.0
10mM dNTPs	0.2
20µM Forward primer	0.2
20µM Reverse primer	0.2
5U/µl Amplitaq Gold® Polymerase	0.08
DNA template	6.3
Total volume	10

Table 3. PCR program.

Step	Temperature	Time
Step 1	94°C	7 min
Step 2	94°C	30 sec
Step 3	51°C	1 min
Step 4	72°C	1 min
Step 5	Go to Step 2	34 times
Step 6	60°C	15 min
Step 7	4°C	∞
Step 8	End	

consisting of 1.5ml Hidi Formamide (Applied Biosystems - Genetic Analysis Grade) and 51µl of GenomeLab™ DNA Size Standard Kit - 400 (Beckman) was prepared, vortexed for at least 90 seconds and aliquots were transferred into 96-well CEQ sample plates. PCR plates were thawed and spun down, then 1.0 or 1.5µl aliquots of PCR product were mixed (and multiplexed in the CEQ plate) with the SLS for loading into instrument. The volume of 1.0 µl was used for blue dye primers because their signal strength was always much higher than that of the other two primers.

Separation buffer (Beckman) was pipetted into a separate 96-well buffer plate and the two plates loaded into the instrument. The PCR products were pulled from the plates through a Separation Gel in the DNA Separation Capillary of the CEQ system. Raw colour-coded fragment data output from the CEQ system were "read" by the CEQ software to generate the allele data for the 15 SSR primers. The output was assessed using the system software to determine which samples needed to be repeated. Regular assessment was necessary to ensure the CEQ system was running optimally. Rainin EDP Electronic pipettes and Gibson multichannel pipettes were crucial for pipetting and a Sigma 4-15C (Qiagen) centrifuge was used for spinning down and mixing samples in 96-well plates.

Results

The complete set of 15 SSR primer pairs was used to amplify DNA from 310 accessions. An additional 71 accessions were run with 9 SSR primer pairs. Peaks were only trusted to be used to "call" the alleles if the standard resembled a typical standard profile. Sometimes atypical standard profiles could be improved by re-analysing the data using system spectra (an option provided by the CEQ analysis software), thus allowing allele peaks to be called. Some samples had to be re-run because they either gave "un-callable" peaks/alleles, no alleles, or had poor standards that could not be improved using the system spectra. Usually problematic samples were re-run through the CEQ system as a first option and if this failed, the PCR was redone. In some cases, there was no improvement in the result obtained. Although the SLS (and DNA standard before addition to Hidi formamide) was vortexed as required and used in the recommended ratio (of Hidi formamide to DNA standard), most of the standard peaks (see

Figure 1) were low in comparison to many of the primer peaks. For some accessions for which results were not obtained, DNA may need to be re-extracted. Samples of ICS and PNG accessions were not included due to being earmarked for SNP analysis. Some samples initially classified as missing were found after an initial search, and updated in the catalogue of all available samples. The profile displayed in Figure 1 is a screen capture of an actual (annotated) profile obtained from three multiplexed primers.

Figure 1. SSR profile: Multiplexed primer pairs P15, P9 and P25.

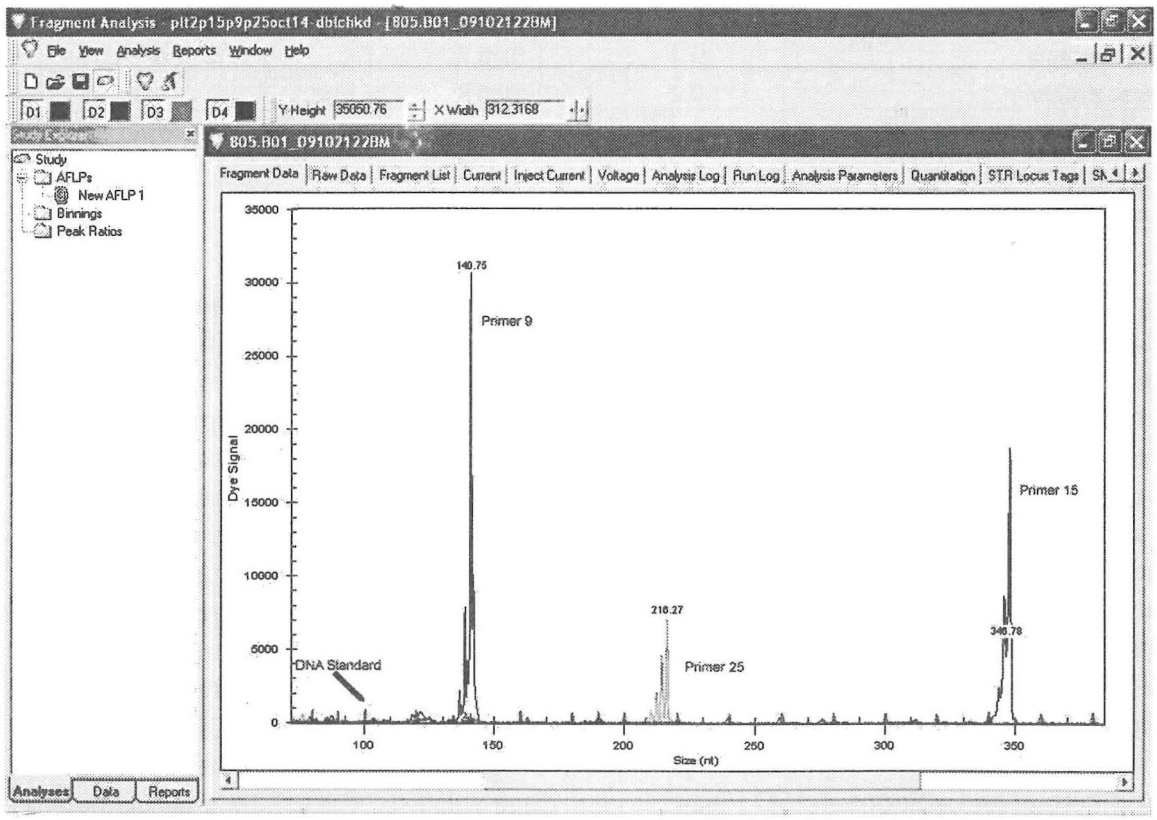

Table 4. Summary of data generated.

No. of primer pairs	No. of accessions with data
1-3	38
4-6	34
7-9	53
10-12	74
13	7
14	23
15	102

The expected success rate for accessions that were difficult to analyse was estimated at <30% (S. Pinney, pers comm) however, when the CEQ system was running well, with a new capillary system, up to 70% success was achieved (allele peaks were "called"). Primer stock quality may have affected success. Table 4 shows the number of accessions for which new data were

generated. Data are available for 331 accessions however not all of the accessions have a complete dataset for all of the SSR primer pairs used. Complete data from all 15 primer pairs is available for 102 accessions.

Conclusions

The assessment of SSR profiles is very time intensive. Although most of the DNA samples were of sufficiently good quality for processing using this method, cacao DNA integrity (especially of samples that were "tried" previously) is one factor that could have adversely affected the output. Also, primer integrity (only old stock was available), DNA size standard quality and capillary system "life" could all potentially have affected the results obtained. Atypical standard profiles that did not improve with the system spectra and occurrences of "pull-up" during the late stages of the capillary "life" led to frequent loss of data, although these occurrences were sometimes ignored in the attempt to prolong the "life" of the capillary. It was not possible to test the efficacy of the primer stock since no new stock was available.

Future Prospects

Data have been compiled in a Microsoft® Excel workbook and this file has been made available for verification work although some accessions are incomplete. Problematic DNA samples will need to be redone, either at a lower concentration, or it may be necessary to re-extract DNA from leaf samples. For those samples tagged as "missing" and that could not be located, DNA aliquots housed at CRU can be used to complete the project. It is possible the project could be completed at Beltsville during another six month period, or if the new sequencing system at UWI has been set up, tested and shown to be functional and the necessary reagents supplied, an attempt could be made to finish the remaining samples with that machine. The one caveat is that with the SSR protocol, completion of the project on the same system is preferred to avoid differences in allele scoring that may exist between the two different though similar types of instruments.

Acknowledgments

Funding for this assignment was provided by the USDA. Special thanks are due to Lyndel Meinhardt (SPCL Director), Dapeng Zhang, SPCL staff: Stephen Pinney, Sue Mischke; Nancy Thomson (BARC administrative assistant), and CRU staff.

References

Boccara, M., Butler, D.R., Motilal, L.A., Sankar, A., Bekele, F.L. (2002) The collaborative CRU/USDA cocoa DNA fingerprinting project: progress report. Pages 24-26 in: *Annual Report for 2001.* St. Augustine, Trinidad: Cocoa Research Unit, the University of the West Indies.

Boccara, M., Johnson, E.S., Butler, D.R., Bekele, F.L., Motilal, L.A., Sankar, A., Sounigo, O., O'Connor, C. and Williams, C. (2001) Sampling strategies for DNA Extraction in the USDA fingerprinting project. Pages 18-22 in: *Annual Report for 2000.* St. Augustine, Trinidad: Cocoa Research Unit, the University of the West Indies.

Saunders, J.A. (2001) USDA DNA Fingerprinting programme for identification of *Theobroma Cacao* accessions. Pages 108-114 in: *Proceedings of the international workshop on new technologies and cocoa breeding*, 16th-17th October 2000, Kota Kinabalu, Sabah, Malaysia. London, UK: INGENIC.

Saunders, J.A., Mischke, S., Leamy, E.A., Hemeida, A.A. (2004) Selection of international molecular standards for DNA fingerprinting of *Theobroma cacao*. *Theoretical and Applied Genetics* **110:** 41-47.