MOLECULAR CHARACTERIZATION OF HIV TYPE 1 DRUG RESISTANCE PATTERNS AMONGST PATIENTS IN BUSIA COUNTY, KENYA

OLIPHER MAKWAGA

A thesis submitted in Partial fulfillment for the requirements of the award of the Doctor of Philosophy in Molecular Biology of the Masinde Muliro University of Science and Technology

NOVEMBER, 2021

DECLARATION

This thesis is my original work and has not been presented elsewhere for a degree or any other award.

Olipher Makwaga Sign: …….. Date: ……...

SBB/H/03/15

CERTIFICATION

The undersigned certify that we have read and hereby recommend for acceptance of Masinde Muliro University of Science and Technology a thesis entitled: 'Molecular Characterization of HIV-1 Drug-Resistance Patterns among Patients in Busia County, Kenya'

Dr. David Mulama, Ph.D Sign: …… Date: …….

Biological Sciences Department,

Masinde Muliro University of Sciences and Technology.

Prof. John Muoma, Ph.D Sign: ……. Date: ……..

Biological Sciences Department,

Masinde Muliro University of Science and Technology.

Prof. Matilu Mwau, Ph.D Sign: …….. Date: ……...

Centre of Infectious and Parasitic-Diseases-Control-Research,

Kenya Medical Research Institute.

COPYRIGHT

DEDICATION

I dedicate this work to my treasured husband, Godfrey Oundo and to my own sons. They all have been my base of enthusiasm and stimulation.

ACKNOWLEDGEMENTS

I fairly intensify my genuine thankfulness to the KEMRI General Director for granting me time that legalized me to commence my Ph. D in Molecular Biology at Masinde Muliro University of Science and Technology. I recognize the Biological Sciences department, Natural Sciences School, for offering me an opportunity to undertake the specified Degree. Earnest indebtedness to my supervisors; Doctor David Mulama, Professor John Muoma, and Professor Matilu Mwau for their mentorship, motivation, and administration they delivered throughout the proposal imprinting, project execution, manuscripts writing, and thesis writing. Heartfelt appreciativeness to the Research grant from KEMRI and the funds of the National Research for backing up of this project. Extraordinary recognitions to Maureen Adhiambo, Humphrey Kimani, and Ferdinard Adungo for their substantial support towards methodological implementation of this work.

ABSTRACT

HIV infection is a communal health tragedy in Kenya causing mortality of 21,300 and morbidity of 1,500,000 cases. The prevalence of HIV infection in Busia County is 7.7% and almost two fold greater than 4.9% of the national. Antiretroviral mitigate HIV-1 infections and viral loads. The coverage of antiretroviral treatment is 95% in Busia County and is greater than 75% of the nationwide amongst adults. The Kenyan HIV type 1 drug resistance prevalence ranged from 3.1% to 23% amongst adults prior to recommendation of the World Health Organization that everybody confirmed HIV positive to be commenced on ARVs. Handful statistics on HIV type1 drug resistance is in existence after the execution of these guidelines. Consequently, the current study was planned to define the HIV-1 drug-resistance amongst adults in County of Busia. This study was conducted for a period of one and half years. Retrospectively, data was obtained from January-June 2019. Cross-section ally, the study was executed from July 2019 to June 2020. Two hundred blood samples were obtained from equivalent number of participants. HIV type 1 viral tallies and RNA transcription to cDNA were executed under real-time PCR. Amplification and confirmation of cDNA were done under Nested-PCR and 1% agarose gel respectively. Cycle-sequencing was executed using 6 primers, Reverse and forward 1, 2, 3. Cycle sequenced cDNA were loaded into the Sanger sequencer to manufacture cDNA sequences in fragments. These were aligned and assembled by Recall software version 3.05 which produced sequences in fasta-format and preliminary reports on drug-resistance. Stanford University, HIV drug-resistance-database was utilized to analyze and deduce HIV-1 drug-resistant-mutations. Subtypes/recombinants were analyzed by jumping profile hidden Markov Model. Sequences were also aligned using CrustalW, and phylogeny tree were drawn by MEGA7. Data analysis employed SPSS version 20, where descriptive statistics was used to compute frequencies/prevalence of various parameters. Chi square test was utilized to compare numerical variables. Out of the total blood samples, 160 (150 ART experienced and 10 inexperienced) were effectively sequenced. Twelve HIV-1 sub-types were identified. A_1 99, 61.9% dominated; D 31, 19.4% followed and C 2, 1.3% was the least for pure isolates. Recombinant subtypes were A_1-B and A_1-D with 7, 4.4% each; A_1-C 6, 3.8%; A_1-G 3, 1.9%; A_1-A_2; A_1-J 1; A_1-K; A_2-D; and B-C had 1, 0.6% each. The general HIV type 1 drug resistant prevalence was 110/150, 73.3%. Single class HIV type 1 drug resistance prevalence were as follows: nucleotide reverse transcriptase inhibitors 88/150, 58.7%; Non-nucleotide reverse transcriptase inhibitors 102/150, 68.0% and protease inhibitors PIs 16/150, 10.7%. The HIV type 1 drug-resistance prevalence was 70.1% for first line and 18.6% second-line therapy. The pattern of HIV type 1 drug resistance prevalence was as follows: dual-class, NRTIs _NNRTIs 46.7%; NRTIs/PIs 0.7% and NNRTIs/PIs 0.7%. Multi-class, NRTIs/NNRTIs/PIs was 8%. None of the ART inexperienced participants had major mutations against available drugs. A total of 464 mutations were identified NRTIs, 207; NNRTIs, 219, and PIs 38. Leading NRTIs mutations were M184V, n=68; K65R, n=23; NNRTIs K103N, n=55; Y181C, n=27; G190A, n=20; PIs were 154V, n=5; M46I, n=5. HIV type 1 subtypes identified presents a contest in the management and control of HIV infection. Consequently, this plea for constant monitoring of the association between the subtypes and drug-resistance if triumphant management of HIV could be achieved. The high HIV type 1 drug-resistance prevalence appeals for the reinforcement of health programs for example creating of testing programs and availing them routinely. The results also boost information to the available literature on HIV-1 drug-resistance.

TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

ABC	Abacavir
AIDS	Acquired-Immunodeficiency-Syndrome
AMPATH	Academic Model Providing Access to Healthcare
ART	Antiretroviral therapy
ARVs	Antiretroviral
ATV/r	Atazanavir
AZT	Azidovodine
CD4	Cluster of differentiation four
cDNA	Complementary Deoxy-Ribonucleic Acid
DNA	Deoxy-ribonucleic acid
dNTP	Deoxynucleoside triphosphate
DTG	Dolutegravir
EDTA	Ethylenediaminetetra acetic acid
EFV	Efevirenz
FDA	Food and drug administration
HIV	Human Immunodeficiency Virus
IQR	Inter-quartile range
KEMRI	Kenya Medical Research Institute
LPV/r	Lopinavir/ritonavir
NASCOP	National AIDS & STD Control Programme
NNRTIs	Non-nucleotide reverse transcriptase inhibitors
NRTIs	Nucleotide reverse transcriptase inhibitors
NVP	Nevirapine

PCR	Polymerase Chain Reaction
PFGE	Pulse field gel electrophoresis
PIs	Protease inhibitors
RNA	Ribonucleic acid
RT	Reverse transcriptase
SERU	Scientific and Ethical Review unit
SPSS	Statistical Package for Social Sciences
STI	Sexual transmitted infections
3TC	Lamivudine
TDF	Tenofovir disoproxil fumarate
TDR	Transmitted drug resistance
UNAIDS	United Nations Programme on HIV and AIDS
USA	United States of America
VCT	Voluntary counselling and Testing
WHO	World Health Organization

DEFINATION OF TERMS

AIDS	The highest advanced stage of infection of the HIV in which the human immune system is debilitated causing Opportunistic Infections.
ART/ARVs	Medicines that are utilized to overwhelm the HIV and stop progression of HIV infection
BLAST	A bioinformatics procedure and database that is utilized to associate sequence data of bases, amino acids, proteins, DNA and RNA.
CRFS	HIV variants that are manufactured during the HIV multiplication within the human body of HIV diseased individuals.
DNA	Hereditary material accountable for development, progression, reproduction and operational of many creatures
HIV	Ribonucleic acid and Retrovirus which infect and deteriorates the immune system of human thereafter killing the CD_4 cells that are purposed to fight it.
NRTI/NNRTI	Classes of HIV type 1 drugs that obstruct the enzymes reverse transcriptase causing the blockage of replication of HIV
PCR	Computerized wet-lab equipment which highpoints to the manufacture frequent copies of Deoxyribose nucleic acid under dissimilar several conditions
PI	HIV medicines that is advanced to mark and stops the activity of the protease enzyme of HIV. This enzyme breaks complex polyproteins into simple proteins which are certainly assembled for fresh daughter virion.
Pol	The highly preserved gene of the HIV that is embraced for reproduction and advancement of the virus.
RNA	Hereditary material that is accountable for instruction, coding, transformation and manifestation of innumerable genes. It transliterated to cDNA during the procedures of polymerase chain reaction

RT	An internal technique which exploits the enzyme of reverse transcriptase to translate cDNA from RNA.
UNAIDS	It is an Organization comprising of 11 united nations. It is referred to as United Nations Joint-Program on HIV/AIDS.
Virologic failure	It is a condition in which the antiviral drugs are unable to work effectively signified through too high viral tallies of ≥ 1000 copies per microliter computed within a three-months pause with adherence delivery between quantification after at least six months of using ART
WHO	It is known as World Health Organization. It regulates health matters globally.

CHAPTER ONE

INTRODUCTION

1.1 Background of the study

Internationally, approximately 75.7m persons have been diseased since the commencement of HIV epidemic with deaths of 32m cases. By 2019, 38m persons were existing with the virus with fresh cases of 1.7m and mortality of 690,000 yearly. The attacks of the epidemic deviates regionally. African states are shattered firm with 4% adults existing with the virus (World Health Organization and United Nations Programme on HIV and AIDS (2019). Kenya scores number 13 with topmost HIV type 1 epidemic worldwide. Approximately, 1.5m individuals are surviving with HIV in Kenya. Amongst grownups, HIV prevalence positions at 4.5% and amongst children, HIV type 1 prevalence positions at 1.3%. Yearly described fresh cases comprises 52,800 people's crossways all ages. About 28,200 HIV interrelated demises were described by 2019. Busia tallies fifth with topmost HIV type 1 prevalence of 38606, 7.7% (Ministry of Health and National AIDS Control Council 2018b) (Appendix 12)

Human-immunodeficiency virus is a retrovirus and lentivirus that cause acquired-immunodeficiency-syndrome (AIDS), a disorder in human in which the immune synchronization flops, occasioning towards lethal unscrupulous infections (Maria *et al.,* 2012). Antiretroviral (ARV) are the medicines that suppress HIV and currently highly-active antiretroviral treatment (HAART) is available for treatment. Three different medicines are pooled in HAART to diminish the pill number and encounter the viral destruction so that to enhance extreme adherence (Thompson *et al.,* 2012) amongst sick individuals. Management with ARV is applauded for entire people diseased of HIV in the

country. Since they aid in the extension of their survival, advance lives, and minimizes HIV transmission (Moraes *et al.,* 2014)**.** The portion of individuals receiving HIV treatment accounts to 1,035,615, 75% amongst adults and 86323, 84% amongst children. About 33654, 95% adults and 3078, 78% children are receiving treatment in Busia County (Ministry of Health and National AIDS Control Council 2018b). This intensification of ART coverage in the overall population is defied with the advancement of HIV type 1 drug-resistance which marques the drugs fail to work powerfully.

Four ARVs classes are available in managing HIV infection in Kenya. These comprises one Integrase-Inhibitor (Quashie *et al.,* 2013), Dolutegravir. Protease inhibitors (PIs) which include Lopinavir/Ritonavir (LPV/r) and Atazanavir/ritonavir (ATV/r). Nucleotide reverse transcriptase inhibitors (NRTIs) such as Tenofovir disoproxil-fumarate (TDF), Lamivudine (3TC), Abacavir (ABC); and Azidovudine (AZT). Non nucleotide-reverse transcriptase-inhibitors (NNRTIs), Nevirapine (NVP) and Efevirenz (EFV) (Ministry of Health 2018). Three of these drugs are assembled into one to make a drug regimen combination so that maximal treatment threshold could be attained to regulate the viral tallies, deterring the commencing of symptoms or progression of AIDS, thus extending survival of HIV infected persons (Julie and Shivaraj 2021).

The assembly of these remedies is as follows: One NRTIs and Lamivudine as an obligatory NRTI remedy and one NNRTI, these are grouped as first line medicines. One NRTI and Lamivudine as a compulsory drug and one PI, these are assembled as second-line medicines. AZT /3TC/NVP or EFV and TDF/3TC/EFV or NVP are celebrated as first streak antiretroviral drug regimen combination for managing HIV among unexperienced adults and adolescents (World Health Organization June 2013a). The authorized

succeeding second line antiretroviral drug regimen combinations are AZT/3TC/LPV/r or ATV/r, TDF/3TC/LPV/r or ATV/r and TDF/3TC/LPV/r or ATV/r (World Health Organization June 2013b).

Nucleotide reverse transcriptase inhibitors are engaged as pro-drugs with phosphorus within a human cell prior to their activity. These drugs in the human cells are activated by Kinases, thereafter exerting their effect through its structure. NRTIs lack 3 primehydroxyl group at the 2 prime deoxyribosyl which enables it to prevents the construction of a 3'-5'-phosphodiester bond in growing DNA chains thus prevent the virus from replication. A chief distinctive of NRTIs drugs is that their amalgamation during RNA dependent DNA synthesis impedes the production of either negative and positive gears of the DNA (Eric and Daria 2012)

Currently, a total of eight NRTIs are in existence: didanosine, abacavir, lamivudine, emtricitabine, zalcitabine, stavudine, zidovudine and Tenofovir disoprovil fumarate, (Eric and Daria 2012).The principal mechanism of achievement is the binding of NNRTI to the reverse transcriptase and the creation of a hydrophobic abridged near active site. This pocket produces a dissimilar spatial prearrangement of the substrate binding location to lessen the over-all polymerase action. By producing unlike outline, DNA production becomes reduced (Eric and Daria 2012). The HIV type 1 protease is an enzyme accountable for the disassociation of the viral gag and gag-pol polyprotein predecessors throughout virion development and act by distracting this disassociation process (Liangqun and Chaoping 2013). Ten PIs are presently approved: lopinavir, amprenavir, darunavir, atazanavir, indinavir, saquinavir, fosamprenavir, nelfinavir, tipranavir, ritonavir.

Nevertheless, two are accessible for consumption in Kenya (Liangqun and Chaoping 2013).

Integrase enzyme associated drug is newly advanced (Quashie *at al.,* 2013). Integrase hurry up response from 3′ prime processing, viral strand and DNA transfer. Entirely all integrase inhibitors spot the strand transmission (McColl and Chen 2010). The discerning consequence on strand transmission is a consequence of certain mechanism of activity in which the inhibitor muddles to the specific complex amongst the integrase and the DNA of the HIV and interconnects with the two important magnesium ion co-factors in the integrase active point and the DNA (DHHS 2021). Therefore, all strands transfer inhibitors encompass two chief components: a hydrophobic set, which connects with the DNA of a virus and the enzyme in the compound and a metal-binding pharmacophore, which confiscates the active location magnesium. Strands transfer Inhibitors are the distinct class of ARV that interconnects with two vibrant elements of the virus, the DNA of a virus and integrase enzyme which forms the substrate for incorporation (Eungi *et al.,* 2018).

The present crystallization of the foamy virus integrase DNA multifaceted with both RAL and EVG (Hare *et al.,* 2010) authenticates the biochemical mechanism and conveys a structural basis for understanding the single degree of action of antiviral agent that has been experimental for elements transfer inhibitors across all HIV (Eungi *et al.,* 2018). Management of HIV infection with NRTIs regularly end into incidences of minimized drug susceptibility HIV type 1 subtypes. Resistance concerning NRTIs is expedited by two machineries: ATP dependent pyro phosphorolysis, where by NRTIs is detached from the 3 prime end of the growing chain and hindrance of termination of chain and amplified discernment between the distinctive deoxy ribonucleotide (Eric and Daria 2012). NRTI

associated mutations materialize in reverse transcriptase. They are characterized as thymidine analog mutations (TAMs). Thymidine analog mutations inspire pyro phosphorolysis. They also are involved in removal of d4T and AZT. TAM amino acid vicissitudes in HIV-1 RT. They encompass two-unrelated paths: the TAM-1 path: T215Y, M41L, L210W and D67N and the TAM-2 path: K70R, D67N, 219E/Q and T215F (Menéndez-Arias 2011).

Another NRTI resistance mechanism involves the walkout of NRTI consolidation inside the growing chain. K65R and M184V/I are example of mutations connected with this mechanism. Treatment with FTC or 3TC lead to advancement of M184V mutation. Treatment with ddC, ABC, ddI, Tenofovir and d4T lead to development of K65R mutation (Menéndez-Arias 2011). In universal, K65R mutation is antagonistic to TAMs and barely advances in persons receiving AZT-based regimen. In the existence of K65R, M184V reinstates Tenofovir vulnerability K65R hardly advances amongst persons receiving Tenofivir who flop FTC or 3TC together M184V (Rhee *et al.*, 2017). In general, NRTI associated mutations diminishes the role of reverse transcriptase and viral replicative suitability. Clinical gain is associated with diminished replicative fitness of NRTI resistant subtypes, however, additional mutations may increase during treatment leading to increased levels of resistance. The forfeiture of replicative fitness may be compensated by other minor mutations (Rhee *et al.*, 2017).

NNRTI associated mutations are resulted by substitutions of amino acid in the its mandatory pocket. Y181C and K103N are the main NNRTI associated mutations (Ming-Tain *et al.*, 2016). Complex forms of NNRTI-resistant-mutations correlate with increase of NRTI associated mutations. There is cross resistance amongst various NNRTIs and

predominantly in the condition of further ancillary mutations (Sluis-Cremer *et al.*, 2015). These mutations hinder HIV reverse transcriptase enzyme from decoding DNA from RNA subsequently obstructing transcription. Presently, four NNRTIs are in existence: delavirdine and nevirapine, etravirine and efavirenz, However, the latter two are obtainable for utilization in Kenya (Eric and Daria 2012). Beside reducing replicative fitness of the virus, sole nucleotide variations with NRTIs cause high-level of resistance (Michele and Robert 2012). A poorer hereditary barrier, negligible impact on replicative fitness, and the slow deterioration of mutations contribute to transmission and steadiness of NNRTI resistant HIV among the individuals. Fascinatingly, mainstream of NNRTI associated mutations developed during treatment with NNRTI are frequently culminate as wild-type sequence (Tebit *et al.,* 2010).

Due to smaller size of 11kDa of protease, it was predicted, resistance against protease inhibitors would be infrequent due to its dynamic utility in life cycle of HIV. However, the protease genome has abundant pliability, with polymorphisms detected in 49 of 99 codons and more than 20 substitutions linked to resistance (Stephanie and Alan 2020). Almost all PIs share comparable biochemical structures and irritable resistance is habitually perceived. Amongst majority of PIs, key resistant mutations clutch closer to the active site of the protease, at sockets positioned by the inhibitor tie position. These amino acid frequent deviances of the amino acids lead to unstable replicative suitability. In addition, deviations positioned esoteric eight chief protease cleavage sites, are linked to PIs resistance (Stephanie and Alan 2020).

Mutations that encrypt resistance against Integrase closely repeatedly design inside the active spot of integrase adjacent to the amino acid remnants that orchestrate the precarious

magnesium co factors. These mutations lead to poor enzymatic performance and replicative capability of HIV. Resistance encoded against Raltegravir is interconnected through three self-regulating pathways of alterations in the integrase gene, as it is defined in key mutations at Q148, Y143 and N155 during clinical trials (Isabelle *et al.,* 2017). These key mutations are detected together with minor mutations for example, N155 are observed with V151L, T97A, G163R, E92Q, and L74M. Substantial cross resistance is observed amongst the strands transfer inhibitors regardless of the prime or minor mutation. Although cross resistance appears to dominate, unrelated drugs encode unrelated (Isabelle *et al.,* 2017).

Tenofovir-disoproxil-fumarate (TDF) obstructs HIV type 1 reverse transcriptase polymerase through binding resentment with the ordinary deoxyadenosine 5'triphosphate and afterwards integration inside the viral DNA resulting to termination of DNA sequence (Sulav *et al.,* 2012). HIV type 1 with condensed vulnerability to tenofovir was observed during culture of cells. The viruses transported K65R replacement within reverse transcriptase and offered a 2 to 4-fold decrease in defenselessness against tenofovir management. Mechanism of action of Tenofovir varies from other nucleoside analogues. It is swallowed and target to achieve a management threshold of 20 and more. The threshold may rise if it is taken with food rich in fat. It has widespread tissue passage, because of its small molecular size and little protein binding. The drug is eliminated unaffected through glomerular filtration and active tubular secretion in urine. Due to this latter feature, alteration of is crucial amongst patients with renal complication. Tenofovir has 10 times half-life more than plasma (José *et al.,* 2008).

Azidovudine (AZT) is another NRTI regimen that the advancement of AIDS amongst HIV infected individuals. It works by suppressing pro-viral DNA. Its active compound, 5-triphosphate holds a strong magnetism for RT that encrypt pro-viral DNA from RNA (Rogers 2008). 5-triphosphate and thymidine triphosphate have similar structures.

Nonetheless, zidovudine 5-triphosphate possess a greater magnetism for RT than thymidine-triphosphate. It has azide as a nitrogen set in place of the ordinary nucleoside-hydroxyl set. Because of this feature RT links 5-triphosphate of zidovudine at the advancing elements of pro viral DNA and DNA synthesis. Duplication is thereafter stopped, since the advancing nucleosides is unable to hold onto the nitrogen set of 5-triphosphate of Azidovudine.

AZidovudine partially halts the activity of some polymerases through a mitochondrial-DNA polymerase due to its discriminating action of RT (Rogers 2008). Therapy with AZT may lead to muscle tissues and heart destruction because the cells of muscles contain a large amounts of mitochondria. In addition, AZT suppress red and white blood cells including bone marrow cells. This leads to fatigue, malaise, anemia etc. As much as AZT blocks the viral replication of HIV. The virus evades and mutate and advancing resistance against the drug. Due to this reason, the drug is combined with other two so that to be able to suppress the virus maximally (Rogers 2008).

Lamivudine is another NRTI that disrupt the synthesis of the DNA of the virus by forming metabolites that compete into viral DNA during phosphorylation. These metabolites block transcription and in turn terminator the synthesis of viral DNA. Because of lack of 3 prime hydroxyl set, pooled analogues of nucleoside blocks the advancement of a 5 prime to 3 prime phosphodiester joining which is vital in the chain of DNA elongation. Lamivudine as an artificial nucleoside analogue, phosphorylate inside the cell to its dynamic 5 prime end triphosphate metabolite. This nucleoside analogue is fixed inside DNA of the virus through the reverse transcriptase resulting into DNA chain ending (Zoe *et al.,* 2006). As much as lamivudine is broadly available, harmless with unresolved toxicity and obligatory

medicine in mixture of all drug regimen. The substandard genetic blockade to resistance, is a paramount weakness that leads to advancement of resistant mutation against 3TC (Diallo *et al.,* 2003). The main mutation M184V against 3TC shrinkages the likelihoods of supplementary impulsive mutagenicity and depresses viral fitness (Wainberg 2004). M184V/I mutations encrypt higher level of resistance against FTC and 3TC, thus decreasing level of resistance to didanosine ddI and ABC and surges susceptibility against AZT, TDF and d4T (Jeannette *et al.,* 2003).

Efavirenz is an artificial purine derivative, parallel to zalcitabine, zidovudine, stavudine, and. Efavirenz was advanced to treat HIV infections amongst individuals flopping treatment with 3TC. Currently, the EFV is administered in combination of other two to mitigate viral threshold. Efavirenz hinders the action of viral-RNA-directed-DNA-polymerase. The mechanism of action of efavirenz depends on intracellular alteration to the vigorous tri-phosphorylated arrangement. The amount of efavirenz phosphorylation deviation is dependent on cell type. The blockage of reverse transcriptase inhibits and manufactures DNA replicas of viral RNA, which, in long run, are indispensable for production of fresh virions. Enzymes inside the cells eventually exterminate the HIV particle that previously had been uncoated during their entry into the host cell (Robertson *et al.,* 2005). Nevirapine is another NNRTI utilized after the immunity has weakened. Even if Nevirapine is utilized well, it is active for a period of time. Nevirapine clinch right to reverse transcriptase and halts the RNA-DNA dependent polymerase activity by consequential disrupting the catalytic location of an enzyme. The activity of NVP does not challenge the template or nucleoside-triphosphates (Nikolenko *et al.,* 2010).

Lopinavir is a protease inhibitor co-administered with ritonavir. Lopinavir-ritonavir has been mixed to manage naive and other PIs experienced individuals. Earlier, 11 mutations

causing substantial variations in vitro susceptibility to lopinavir were discovered in the protease gene of HIV. Consequently these 11 mutations forms the foundation of mutation score of lopinavir (Robyn *et al.,* 2011). On the other hand, a former study conducted in China, proved that, lopinavir based drug was tolerated amongst HIV-1 infected drug-naïve individuals (Bin *et al.,* 2019). Atazanavir is an azapeptide HIV-1 PI that prevents the formation of mature virions through the potent and selective inhibition of viral Gag and Gag–Pol polyprotein processing in HIV-1-infected cells. Although atazanavir can be administered alone to treatment-naive patients who cannot tolerate ritonavir, it is preferably co-administered with low-dose ritonavir to optimize its pharmacokinetic parameters and efficacy (Chad *et al.,* 2011).

HIV type 1 drug resistance is defined as a syndrome in which the drugs are incompetent to kill and terminate the virus from the human body system. HIV type 1 drug resistance is triggered by nonconformities in the genome of the virus compelling them to progress mutations. Mutation is discreetly a modification in the genome of alive creature (Cuevas *et al.,* 2015). HIV is a retrovirus RNA virus, shortage of proof-correcting activity throughout reverse transcription of complementary DNA from viral RNA. The virus' Reverse transcriptase enzyme is shortened with the proof-reading competences which is intrinsic in cellular polymerases, connoting that, its duplication of the gene of HIV is exceedingly error prone. Mutations that occur as a result of uncorrected Reverse transcriptase action present in the genome of HIV at a degree of about 1 per 1500 to 4000 nucleotides in every replicating cycle making mutated fresh viruses unlike from the mother' virions (Cuevas *et al.,* 2015).

Historically studies documented some mutations encrypting resistance against various sorts of therapeutically drugs in Kenya (Budambula *et al.,* 2015, Hassan *et al.,* 2014, Kinyua *et al.,* 2018, Kiptoo *et al.,* 2013, Onywere *et al.,* 2017, Sampathkumar *et al.,* 2014) A study implemented in the year 2004 amongst drug- inexperienced and experienced folks in Southern Korea reported that, 50% of HIV infected individuals with virologic failure harbored HIV-1 viruses with numerous resistant mutations and 10% multiclass HIV-1 drug-resistance (Min *et al.,* 2013). Approximately 3.1 to 23.1% of HIV-1 drug-resistance prevalence was reported in Kenya between the years 2005 and 2010. Around 6.4% and 3.1% drug experienced and naïve participants respectively resisted PIs drugs in a study executed historically in coastal area of Kenya were (Budambula *et al.,* 2015) Implication that, the utilization of drugs may lead to the advancement of drug resistance.

One past study executed at the National hospital of Kenyatta reported 16.9% and 21.5% of the samples had HIV type 1 encrypting resistance against NRTIs and NNRTIs respectively (Kinyua *et al.,* 2018). Additional study which sequenced the entire genome of the virus amongst drug naïve participants reported that, 22% had mutations encrypting resistance against PIs (Sampathkumar *et al.,* 2014). Connoting that the individuals might have acquired HIV strains which were already resistant to drugs. A further study executed 10 years back documented that, 13.8% of HIV type 1 infected drug users had mutations encrypting resistance against NRTIs_NNRTIs.

HIV type 1 drug resistance is accruing in Kenya. There exist numerous influences that are contributing to this upwelling including assortment of resistant mutations in the existence of drugs (Heestermans *et al.,* 2016, Azia *et al.,* 2016, Burns *et al.,* 2019). Presently everybody diseased of HIV is commenced on treatment with ARVs heedlessly of his/her

CD4 scores comparable it used to be occasionally back. Up scaling of counties and country's ART coverage exists. Busia County, tops with the uppermost (Ministry of Health and National AIDS Control Council, 2018b). It is important for someone to be checked of HIV-1 drug resistance before he/she is switched onto another drug or commenced on drugs. However, routine laboratory HIV type 1 drug resistance testing is not available (Ministry of Health and National AIDS Control Council, 2018b). Additionally, there is upsurge in variance of HIV type 1 subtypes of which past studies documented their relationship with drug resistance (Kityo *et al.,* 2017).

Consequently, the present study aimed at classifying the current mutations encrypting resistance against available ARVs classes. This info will advance the information on the evolutionary pattern of HIV type 1 in Kenya and entice an amendment of treatment alternatives. Much info on the snowballing drug resistance are based on transmitted drug-resistance and in main towns (Aghokeng *et al.,* 2011, Hamers *et al.,* 2011, Hassan *et al.,* 2013, Ndembi *et al.,* 2011). Equally, accounts on the overall HIV type 1 drug resistance is wanting primarily in remote places which have had longer period of utilizing ARV. Furthermore, the info is inadequate after the effecting of the World Health Organization policies that everyone diseased of HIV to commence utilizing ARVs promptly after confirming positive heedlessly of the CD4 totals (Ministry of Health 2018).

In totaling, the past documented study which reported on HIV type 1 drug resistance in County of Busia was executed almost a decade years back. This study considers kids of five years and below to describe HIV type 1 variants and acquired HIV type 1 drug-resistant mutations (Lel *et al.,* 2014). To supplement to these, the up-to-date and current statistics amongst older folks is inadequate simply due to partial studies executed amongst these

individuals of age set and particularly in Busia County with HIV rampant. Hence, the present study was organized to describe HIV type 1 drug resistance amongst grownups living in Busia County (Min *et al.,* 2013). This info will support in the accounting of the gaps in the setting of health systems amongst overall HIV type 1 diseased people in remote areas, which contain the majority of the infected with HIV type 1 specifically in Busia County.

Amongst individuals taking ART therapy, HIV type 1 drug resistance stands an unwarrantable subject because some people who commenced treatment during the start and midst of 1990's have progressed with mutations encrypting resistance against several classes of drug regimens. To supplement to this, some already resistant variants of HIV are acquired by, furthermore, people on ART therapy remain to increase internationally. Disappointing, increasing virus due to the development of drug resistant mutations creel the victory of managing HIV infection (Pleuni 2013).

HIV type 1 drug resistance tests may guide health officers on the selection of therapeutical options for their clients. The tests are not habitually accessible in Kenya. Inadequate labs execute these tests precisely for research activities. Two approaches are in existence in assessing HIV drug resistance. These include, phenotypic and genotypic drug methods. Phenotypic approach determines the concentration of an antiretroviral agent at which inhibition takes place (Günthard *et al.,* 2019) and entails co cultivation of the individuals' marginal blood mono-nuclear cells together with sero negative contributor to yield stocks of viruses, titrating the stock of the virus to describe the infection of the virus, utilizing a standardized inoculation to taint cultures at changing concentrations of ARV agent and calculating IC_{50} on the basis of a grade of contamination. The disadvantages of this

methods include labor intensive and time consuming thus takes at least 6 weeks after sample gathering up to manufacturing of results. Moreover, the method selects some variants in-vitro. The does not directly estimate the virus in the plasma (Günthard *et al.,* 2019).

Genotypic test defines various event of mutations which are patented against encrypting abridged drug-susceptibility. Genotypic assays are set to determine the sequences of nucleotide which encrypt phenotypic-resistance. Dideoxynucleotide sequencing is a method utilized to sequence RT and PI genes of HIV. The method depend on sequencers (Günthard *et al.,* 2019). Sequence alignment, sequences editing, detection of mutation and interpretation require additional software. Due to automation of these processes, the time of processing the samples and production of results have been reduced (Olipher *et al.,* 2020). Additional sequencing tests which involves sequencing by hybridizing oligo-nucleoside probes are available. Most of the RT sequence and whole protease are determined by sequencing to reduced great density ranges of oligo-nucleoside probes. This is followed by assembling of fragmented sequences and determination of the mutation (Gene Chip; Affymetrix) (Radmila *et al.,* 2017). LiPA; Inno-Genetics is hybridization based-line probe procedure which sequence info inadequately which speedily and alongside controls the manifestation of pre-selected HIV drug-resistance mutations in most of the codons (Radmila *et al.,* 2017).

Abundant parameters link to all available drug-resistance assay. All assays require plasma consisting ≥ 500 to 1000cp/ml. Plasma sample with little copies of RNA do not produce dependable results. Detection of virus amongst earlier treated individuals with different ARVs may be difficult because during the initial weeks of termination of a drug, the leading

HIV-1 population may change from HIV with resistant mutations which are specific to drugs and looks like susceptible to the reserved drug (Hamlyn *et al.,* 2012). However, resistance will promptly resume again when the drug is restored. Many genotypic assays identify marginal populaces at stages as minute as about 20%. However, LiPA identify at little as about 4% (Radmila *et al.*, 2017).

The integrase inhibitor class of ARV is the latest to be legalized for management of HIV diseased folks. The drugs hinder the secondary stage that speed up by integrase enzyme by clinching to active site of the enzyme. Integrase dislodges the 3 prime end of viral-DNA from the site that is active and it consequently detach the di-Valente ion that is vital for enzymatic action of integrase (Margolis *et al.,* 2015). Presently, three Integrase drugs have been approved for management HIV infection: Dolutegravir, raltegravir, elvitegravir (Quashie *et al.*, 2013). Other two integrase drugs (Bictegravir and Cabotegravir) are under investigation (Margolis *et al.,* 2015, Tsiang *et al.,* 2016).

Virologic failure is incapability of any antiviral drug to kill or suppress viral scores to unnoticeable level. Factors causing virological failure have been reported in the past studies. Adherence is among the frontier factor causing virological failure. This factor is linked to long distances from families to the health facilities. It is also associated with expensive cost of transport of which participants have to incur while visiting the health amenities. The relationship of health care provider and the patients matters a lot. Family commitments, stigma, long waiting periods at the facilities, work, forgetfulness directly influence virological failure (Heestermans *et al.,* 2016). Many past studies have stated that virological failure is caused by adherence (Azia *et al.,* 2016). HIV-1 drug resistance due to

selection of mutations in presence of drugs is another factor contributing to virological failure.

One study showed that, 40% of the participants in developed countries had persistent viral load of above 1000 copies/ml for over 12 months on therapy and harbored major mutations conferring resistance against the administered antiretroviral (Burns *et al.,* 2019). Consequently, some selected mutations during treatment encode for considerable phenotypic-resistance by themselves, while further mutations upsurge resistance when other mutations are present or compensate for the reduced replicative action that can be linked to drug-resistance (Robert *et al.,* 2001).

This entails alteration in genomic structure of a virus leading to advancement of mutations. These mutations cause alterations of some proteins which involves in the replication of the virus. All RT host mutations in the HIV genome thus leading to advancement of resistance to antiretroviral. Subtypes of HIV advance resistance if they are inadequately killed or suppressed by the antiviral agents which are typically advanced to destroy them (Cuevas *et al.,* 2015). HIV duplication is a vital backer towards the rapid and widespread advancement of resistance. Principally, HIV RT enzyme is tremendously non-selective during transcription of DNA from RNA. HIV RT creates approximation of one error in each HIV genomic material per one round of replication (Cuevas *et al.,* 2015). Most of these mistakes are due to base insertions, substitutions and duplications. Furthermore, HIV has prodigious replication rate in that billion new virions are produced daily amongst untreated individuals (Cuevas *et al.,* 2015).

HIV reserve the infection at a steady level by infecting new cells at a very prodigious rate. This feature makes the HIV to replicate and proliferate very fast leading to many different

variants. Patients infected with HIV may harbor numerous different virus in their system. These differences may have dissimilar responses to antiretroviral agents (Eric and Daria 2012). This scenario confuses the advancement of drugs for managing the infection. Another feature that results to the development of HIV drug-resistance encompasses inappropriate selection of antiretroviral, reduced patient acquiescence and sub therapeutic blood intensities of antiretroviral (Eric and Daria 2012). Follow ups are needed to remind patients to take their medications without fail. HIV type 1 is principally divided into incredible genetic inconsistency because of unusual rates of mutation and mixing. This has driven HIV cataloguing in the three tenuously reliable clusters; Principal set, an outlier set, and non-O-non M set (Joris *et al.,* 2019). The M set commands the AIDS epidemic and is partitioned into 12 varied genetics, entitled as sub-types and 33 circulating recombinant subtypes (Marta *et al.,* 2020).

Third world countries are dominated with A and D subtypes (Marta *et al.*, 2020). Subtype B dominates in Brazil (Tiago *et al.,* 2021); Europe (Beloukas *et al.,* 2016); Thailand ; USA and Australia (Alison *et al.,* 2017). Sub-type C is necessarily circulated in Ethiopia; India and South Africa (Matthew *et al.,* 2020). F circulates in Europe; Central Africa and Eastern Africa (Beloukas *et al.,* 2016). Circulating recombinant A_E is distributed in eastern southern of Asia including China (Minjie *et al.,* 2017). Past studies reported that popular of cases of HIV type 1 in Kenya are sub-type A that is extremely deviating comparable with other sub-types (Kageha *et al.,* 2012, Kinyua *et al.,* 2018**).** This is since sub-type A was presented in Kenya far earlier and is exceedingly communicable likened other sub-types (Khamadi *et al.,* 2005).

Other archaeologically studies demonstrated elevated sub-types C and D prevalence (Adungo *et al.,* 2014, Kageha *et al.,* 2012, Oyaro *et al.,* 2011). Sub-types C and D dominate in Tanzania (Joris 2012) and Uganda (Susanna *et al.,* 2020) respectively. These county edges Kenya, therefore the existence of these two sub-types in Kenya could be due to its introduction because of cross boundary migration persons from Uganda and Tanzania into Kenya. As significant as stated above ascertainment occurred concerning HIV-1 subtypes in Kenya, all foreseeable investigations were executed numerous years back. Consequently, the current study purposed to describe the present mingling HIV-1 subtypes amongst participants attending selected health facilities in the County Busia, Kenya. Other previous studies documented HIV type 1 recombinant sub-types in different counties of Kenya including Busia (Adungo *et al.,* 2014, Khoja *et al.,* 2008, Otecko *et al.,* 2016, Oyaro *et al.,* 2011)

HIV type 1 drug resistance is a syndrome in which the drugs are incompetent to suppress the virus and the virus continue to replicate in the presence of drugs. HIV type 1 drug-resistance is commenced with some adjustments in the genomic material compelling the virus to develop mutations. Mutations is fundamentally a modification in the genomic material of any organism. HIV type 1 is retrovirus RNA and lacks proof-reading action while transcribing complementary DNA from RNA. This proof reading incompetency is an inherent condition in the cellular polymerases leading to errors during duplication. Mutations that occur due to un-corrected RT action are found in the HIV genomic material at a degree of approximately one in every 1500 to 4000 nucleosides in each replication sequence (Cuevas *et al.,* 2015). Studies executed out historically described encrypting resistance against classes of drugs (Hassan *et al.,* 2014, Kiptoo *et al.,* 2013). Accordingly,

the current study was aimed at classifying supplementary mutations encoding resistance against available drug regimen. The outcomes of this study will strengthen the knowledge of evolutionary group of HIV type 1virus and captivate an adjustment of treatment selections in Kenya.

Substantial info about the rising drug resistance are transmitted type of drug-resistance (Aghokeng *et al.,* 2011, Hamers *et al.,* 2011, Hassan *et al.,* 2013, Ndembi *et al.,* 2011). Similarly, documentation of the general HIV type 1 drug resistance is inadequate particularly in remote areas where ARVs have been utilized for a longer duration. Moreover, the info is scarce after the implementation of the World Health Organization instructions that everyone infected with HIV to initiate taking ARVs promptly after confirming HIV positive heedlessly of the CD_4 totals (Ministry of Health 2018).

With accounting, the past study which described HIV type 1 drug resistance in the County of Busia was executed more than half a decade ago and focused amongst children of less than 5 years old to describe HIV type 1 sub-types and acquired drug resistant mutations (Lel *et al.,* 2014). Equally, the newest and latest data amongst people who are older than the earlier studied individuals is inadequate merely since inadequate studies amongst these persons have been executed. Therefore, the current study was deliberated to describe HIV type 1 drug resistance amongst grownups living with HIV in the County of Busia (Min *et al.,* 2013). This information will support in the fixing of the gaps in the improving of health systems amongst the overall HIV type 1 diseased individuals in remote locales, which comprehend the majority of those infected primarily in the County of Busia.

Resistance has remained a calamitous clash for the subsequent reasons: Some sub-types of virus already carrying resistant mutations are being conveyed. Additionally, persons who

are treated remain to surge internationally comprising Kenyan population. Furthermore, some folks who initiated therapy in 1990's have progressed viruses with mutations encrypting resistance against various drug regimen. Moreover, they are initiated or swapped to additional drugs in absence of drug resistance tests. Inadequate virus suppression due to the progressive resistant mutations impede the victory of HIV management (Olipher *et al.,* 2020).

A study executed in 2004 amongst drug experienced/unexperienced individuals in South Korea reported that a half of participants showing virological failure harbored mutated viruses and 10% multi-class HIV type 1 drug resistance (Min *et al.,* 2013). Approximately 3.1 to 23.1% prevalence of HIV type 1 drug resistance were reported in studies executed in more than a decade years ago in Kenya. Therefore, the current study was intended to describe the existing prevalence of HIV type 1 drug resistance amongst partakers of ARVs.

1.2 Statement of the Problem

Virologic failure due to HIV type 1 drug resistant mutations amongst HIV type 1 diseased individuals have been reported locally (Olipher *et al.,* 2021). This occurred due to unpredictability of the HIV type 1 genome that consist error prone proof reading RT enzyme during transcription of cDNA from RNA, mixing through duplication, sudden turnover of HIV type 1 in vivo and host selection of pressure on immunity (Désiré *et al.,* 2018). The advancement of drug resistant mutations poses a risk to the whole treatment efforts geared towards management of viral infection. Since the commencement of usage of ARVs, there is no up-to-date info in Kenya on the of HIV type 1 drug resistance prevalence in the studied region. Hence outcomes from the current study will boost the prevailing knowledge on HIV type 1 drug-resistance.

1.3 Justification of the Study

By 2018, Kenya nation was ranked number 13 in the world with the highest HIV infection (Ministry of Health and National AIDS Control Council 2018b)**.** Anti HIV drugs support people living with HIV extent their lives, live healthier and minimizes the danger of spreading HIV. Nevertheless, drug resistance caused HIV-1 accumulate major public healthiness issues in the entire community and intimidates management of infection due to HIV. Historical studies executed in approximately a decade years back documented mingling HIV type 1 subtypes and drug-resistance in unexperienced grown-ups (Hassan *et al.,* 2013, Kantor *et al.,* 2014, Kiptoo *et al.,* 2013, Zeh *et al.,* 2011)**.** Conversely, inadequate up-to-date such data exists. Therefore, it was essential for such current study to describe the mingling variants of HIV-1 and mutations amongst HIV diseased individuals in the County of Busia. So that appropriate management action could be adopted for supporting these group of individuals.

Previous studies on the HIV type 1 drug resistance patterns described mutations encoding resistance against multiclass, dual class and single class drugs amongst folks taking both lines of drug regimen combination (first and second) (Budambula *et al.,* 2015, Hassan *et al.,* 2014, Kiptoo *et al.,* 2013, Lihana *et al.,* 2009, Onywera *et al.,* 2017, Sampathkumar *et al.,* 2014, Sigaloff *et al.,* 2012, Steegen *et al.,* 2009). Multiclass drug associated mutations (NNRTIs, NRTIs and PIs) were determined amongst grownups in Nairobi (Koigi *et al.,* 2014), as well as single class drugs, PIs associated mutation amongst grownups injecting drug users in the coastal setting of Kenya (Budambula *et al.,* 2015)**.** Inadequate information of such kind in other locales of Kenya including Busia where HIV is rampant exists. Therefore, the current study will describe the of HIV-1 drug resistance pattern.

1.4 Research Questions

1. What is the prevalence of subtypes of HIV type 1 in the County of Busia?
2. What is the HIV type 1 drug-resistance prevalence amongst individuals taking ARV treatment in the County of Busia?
3. Which mutations and patterns of HIV type 1 drug resistance are existing among individuals taking ARV treatment in the County of Busia, Kenya?

1.5 Objectives

1.5.1 General Objectives

To determine the HIV-type 1 subtypes, prevalence and characterize HIV-type 1 drug resistance mutations amongst patients in Busia County, Kenya

1.5.2 Specific Objectives

1. To describe HIV type 1 subtypes prevalence among participants in the county of Busia, Kenya
2. To describe HIV type 1 drug resistance prevalence amongst individuals utilizing ARV therapy in the County of Busia, Kenya
3. To describe HIV type 1 drug resistant mutations and patterns among individuals utilizing ARVs therapy in the County of Busia, Kenya

1.6 Significance of the Study

The outcomes gathered from the present study will enhance proof of earlier existing data of related studies about the mingling HIV type 1 sub-types/drug-resistance; Info out of the current study will consequently explain verdict implementation regarding management ranges amongst diseased people; Data of the current study appeals for health programs strengthening in Busia County including the Country at large.

1.7 Limitations of the Study

The Reverse Transcriptase and Protease of the pol gene were amplified and sequenced which might have underestimated the frequencies of mutations. Therefore, in such circumstances I endorse all the enzymes of the pol gene to be sequenced upon availability of kits and protocols that can embrace amplification of all the three enzymes (Reverse Transcriptase, Protease, Integrase), now that recently, integrase drugs have been initiated for use amongst Kenyan population. However, the sequencing of the two enzymes did not have any influence on the subtypes in that the two genes are still conserved regions of the HIV-1 virus. To add on these, studies conducted in more than 2 decades ago on the identification of the subtypes and drug resistance have been sequencing these two genes of the HIV-1.

CHAPTER TWO

LITERATURE REVIEW

2.1 Extent of HIV type 1 infection and Coverage of ART

Human Immunodeficiency virus has resulted into infection of nearly 75.7m individuals worldwide since the start of the virus in early 1980's with approximately 32m deaths. By 2019, occurrence of 38m cases were surviving with HIV; 1.7 million new cases and 690000 deaths were reported (World Health Organization and United Nations Programme on HIV and AIDS 2019). Majority of the effects of the epidemic differs as per the region. Third world nations are afflicted the most whereby one in 25 grownups exists with virus (World Health Organization and United Nations Programme on HIV and AIDS 2019). Kenyan nation is enumerated number 13 out of the total global nations bearing the uppermost endemic of 1.5m cases with the infection, thus 1388200 grownups and 105,200 children are infected (Ministry of Health 2020, Ministry of Health and National AIDS Control Council 2018b). The prevalence of HIV among the grownups is 4.5% with 52,800 newly infected cases through the whole age sets. Approximately, 28,200 interrelated demises dues to AIDS were reported during the year 2019. Over ally, the County of Busia is enumerated 5th with the highest HIV-prevalence, thus 38606, 7.7% of the population is infected (Ministry of Health and National AIDS Control Council 2018b)

Antiretroviral treatment/therapy is recommended as a standard management practice amongst HIV infected individuals due to their activity in mitigating viral loads and boasts immunity against other opportunistic diseases (Ministry of Health and National AIDS Control Council 2018b). About 26m cases are utilizing antiretroviral globally. Approximately, 1035615, 75% of infected grownups and 86323, 84% children are utilizing

antiretroviral. Hence, the County of Busia enumerates additional coverage of grownups who are receiving antiretroviral therapy thus 33654, 95% and children 3078, 78% than that recorded by the Nation (Ministry of Health 2020, Ministry of Health and National AIDS Control Council 2018b)**.** This is due to many newly diseased individuals are initiated to take antiretroviral without checking levels of their CD_4 cells that play a role in mounting an immunity of destroying HIV. HIV due to HIV-1 has persisted to rise in the country; for example, further 44000 freshly infected persons were reported in the year 2017 (United Nations Programme on HIV and AIDS 2018)

2.2 HIV-1 Subtypes

Various epidemiological investigation help to monitor the extent of global HIV transmission (Jones *et al.,* 2019). Through cumulative of HIV prevalence, distribution of sub-types is ghettoized to a greater geographical ranges. The extreme genetic variation among the HIV-1 is found in regions where the endemic is stalest such as the third world nations where highest frequency of the HIV type 1 subtypes and numerous of the mingling recombinant have been reported (Marta *et al.,* 2020).

Some studies executed in parts of Uganda reported subtypes A, D and their recombinant variants to be predominant (Deogratius *et al.,* 2020). Numerous historical genotypic studies in Tanzania reported the dominancy of HIV-1 subtypes C_D, A_2, A_1, D, C, G to co-circulate among HIV-1 infected persons (Erik *et al.,* 2017, Ireen *et al.,* 2012). Additional study executed in more than a half a decade years back documented subtypes D, A1, C, A_2, and G circulation along the borders of Kenya and Uganda (Adungo *et al.,* 2014). The study region is located at the borders of Kenya and Uganda with two border posts which aid as entrance paths for travelers to and fro Congo, Rwanda, Uganda, Tanzania and

Somalia. With the help of this study region, frequent fresh HIV variants are propelled into the Kenyan populace. The introduced different HIV-1 variants could carry mutations that encode resistance to available drugs creating a hash condition for managing HIV contagion. Additional study conducted within major cities of Kenya described a majority of the variants were pure A subtype and inter-subtype A recombinant variant (Kamini *et al.*, 2017). Another previous study showed that subtypes C, D, A1, dominate in Central areas of Kenya (Kageha *et al.*, 2012)**.** Though subtype A seems to dictates utmost areas of Kenya Nation, the prevalence of other inter-subtypes, recombinant viruses and pure subtypes are equally accumulating (Kamini *et al.*, 2017). Another further study conducted in Nairobi had similarly reported raised existence of subtypes D and C (Veronica *et al.*, 2012) in Kenya. Another supplementary study on HIV-1 drug resistance amongst children of 5 and below years in Busia county reported the existence of subtypes C, A, D and A_1C in circulation (Lel *et al.*, 2014). Therefore, there was need for this study so as to identify the current HIV-1 subtypes in circulation along this viable border and especially amongst adults.

International Map Presenting HIV type 1 Subtypes Distribution

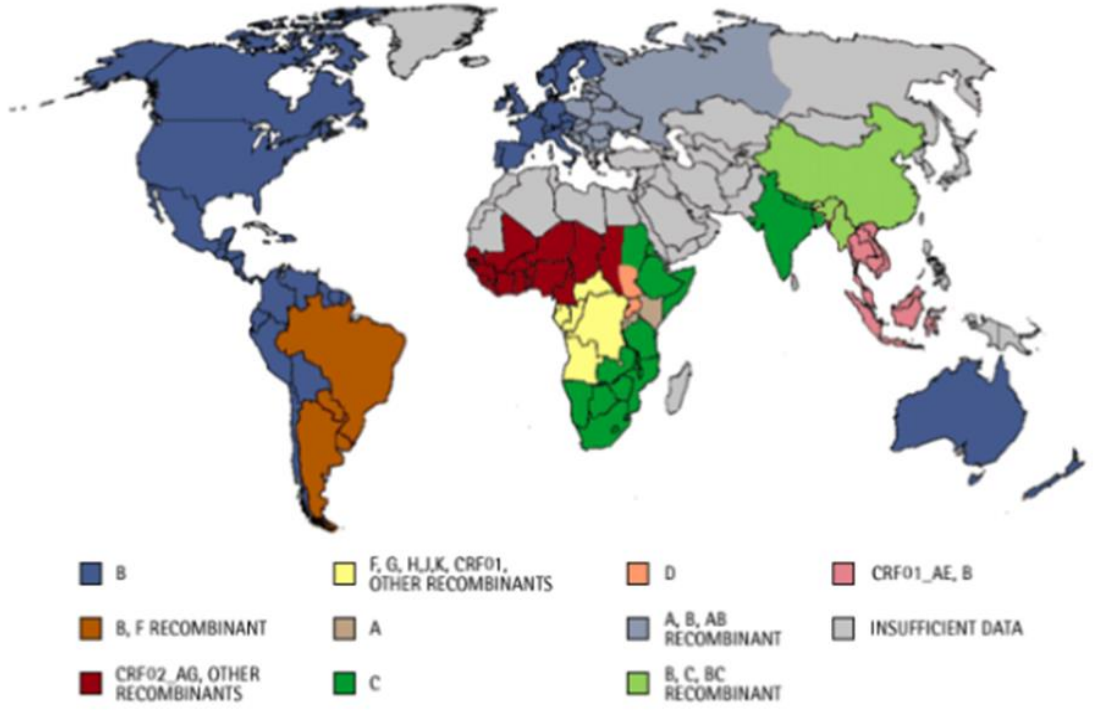


Figure 1:- International Dispersal of HIV type 1 Subtypes.

Source: - HIV type 1 International dispersal at pbs. org/wgbn /pages /frontline /aids/ atlas/clade. html

2.3 Extent of HIV type 1 Drug Resistance

Human Immunodeficiency Virus type 1 drug resistance is among the leading communal health issue in the overall populace and has become a hazard in management of HIV infection. Resistant HIV-1 forms advanced while utilizing ART are prone to blowout to others. Drug resistant strains have been defined in ART unexperienced people. For example, a logical assessment showed that the overall drug resistance HIV type 1 prevalence spread was 4.2% in Asia, 12.9% amongst North Americans, 10.9% amongst Europeans, 6.3% amongst Latin Americans, 4.7% among Africans (Frentz *et al.*, 2012).

Consequently, the larger prevalence of drug resistant HIV spread was described amongst individuals in regions with advanced coverage ART uptake and commonly in industrialized countries. It is substantial to understand that, along with ART scale up, the acquired HIV type1 drug-resistance prevalence raised to 5.3% in late 2003 from 2.8% in early 2001 in sub-Saharan nations (Frentz *et al.,* 2012). Contrary in 2011, approximately, 5% of acquired drug resistance prevalence was reported amongst third world nations (Bussmann *et al.,* 2011, Parboosing *et al.,* 2011)

Further studies in sub-Saharan countries reported increased drug resistance prevalence in some settings. In Uganda, transmitted-drug resistance prevalence rose to 8.6% from 0% in 2009 to 2010 and 2006–2007 respectively (Ndembi *et al.,* 2011). Another cohort study carried out amongst HIV exposed populations in Zambia reported raised transmitted drug resistance prevalence of 16% from 0% in 2005 (Price *et al.,* 2011). In Cameroon, a static raised drug-resistance prevalence of 12.3%, 2007 form 0%, 1996–1999 was detected. Strangely, the similar study described drug-resistance prevalence of 4.8% in 2006–2007 in rural settings of the same country (Hamers *et al.,* 2011). In a cross sectional study conducted in Nairobi, Kenya in 2005, 7.5% of participants were reported to have acquired HIV type 1 drug resistance (Hassan *et al.,* 2013).

Approximately 3.1% prevalence of acquired drug-resistance was reported by an International AIDS Vaccine Initiative (IAV) (Price *et al.,* 2011) at the start of early infection cohort between 2006 and 2009 (Price *et al.,* 2011). Another cross sectional study executed by President's Emergency Plan for AIDS Relief (PASSER) group between 2007 and 2009, reported transmitted drug resistance prevalence of 4.5% in Mombasa and 4.9% in Nairobi (Hamers *et al.,* 2011). Additionally, a cross sectional study reported 13.2%

transmitted drug resistance prevalence amongst ARV unexperienced grownups visiting four Voluntary counselling and Testing centers (VCT) in Mombasa between 2009 and 2010 (Sigaloff *et al.,* 2012)**.** A Surveillance investigation conducted during the year of 2012 in the remote areas of western places of Kenya documented a 9.2% prevalence of acquired drug-resistance (Onywera *et al.,* 2017).

Several studies implemented in more than a decade back in Kenya have been on transmitted form of drug resistance (Hassan *et al.,* 2013, Kantor *et al.,* 2014, Kiptoo *et al.,* 2013, Zeh *et al.,* 2011). These studies concentrated on the antiretroviral unexperienced antenatal persons and antiretroviral unexperienced grownups (Hassan *et al.,* 2013). About 23.1% HIV-1 drug resistance prevalence was reported in a retrospective cohort study whose specimen were sourced in more than a decade years back among individuals on first line therapy at National health facility of Kenyatta (Kinyua *et al.,* 2018). Thus, a larger number of studies in Kenya have documented acquired and transmitted type of drug-resistance and especially in urban settings (Aghokeng *et al.,* 2011, Hamers *et al.,* 2011, Hassan *et al.,* 2013, Ndembi *et al.,* 2011), or among high endangered populations (Price *et al.,* 2011)**.**

Another study implemented in 2012 and 2013 reported that, 31.3% and 40.9% of patients had failed on second and first line drug regimen treatment respectively (Koigi *et al.,* 2014). Up-to-date scanty studies have been executed to describe drug resistance prevalence in rural settings with a longer experience of utilizing ARV. Furthermore, all the studies implemented in Kenya were comprehended earlier prior to the World Health Organization rules that all individuals infected with HIV and confirmed positive to be initiated on drugs as soon as possible. Consequently, it was commendable to describe the present prevalence

of HIV-1 drug-resistance amongst participants who are utilizing HIV anti-drugs afar these WHO recommendations

2.4 Pattern of HIV-1 Drug Resistance

Info concerning HIV type 1 drug-resistance patterns amongst HIV exposed individuals is vital in development of new operative drugs. Several ARVs have been permitted in managing HIV-1 disease (World Health Organization 2018). The drugs include nucleotide and nucleoside reverse transcriptase inhibitors, protease inhibitors, nonnucleoside RT inhibitors, and fusion inhibitor. The nucleoside analogs are zidovudine, didanosine, zalcitabine, stavudine, lamivudine, abacavir and emtricitabine. Protease Inhibitors are indinavir, amprenavir, ritonavir, nelfinavir, saquinavir, Atazanavir and lopinavir which are combined with ritonavir (World Health Organization 2018).

Historical researches have reported that the prevalent cases are living with HIV harboring mutation encrypting resistant against NRTIs tracked by NNRTs and infrequent cases are resistant against PIs. Studies carried out concerning HIV-1 drug resistance patterns in Kenya have reported mutations encrypting resistant against PIs, NNRTs and NRTIs, and amongst grown-ups (Budambula *et al.,* 2015; Koigi *et al.,* 2014; Onywera *et al.,* 2017). Thus approximately 40.9% of participants failing on first line therapy harbored viruses with mutations encrypting drug resistance against NRTIs. Similarly, 30% participants failing second line treatment harbored viruses with NNRTIs linked mutations (Koigi *et al.,* 2014). Mutations encoding resistant against PIs were reported amongst drug injecting grownups along the Coastal areas of Kenya (Budambula *et al.,* 2015). Scarcity of alike info exists in some region of Kenya consisting the county of Busia where HIV remains rampant.

Therefore, the current study was organized to describe HIV-1 drug resistance pattern amongst vulnerable HIV-1 infected participants in the County of Busia.

2.5 HIV-1 Antiretroviral Treatment

Anti-HIV drugs are progressively available in sub-Saharan nations including Kenya (United Nations Programme on HIV and AIDS 2018). Info about the background of HIV-1 disease and tallying of RNA is similarly vital: These tasks are superlative accomplished by usage of operative ART to exceptionally block HIV reproduction so that viral loads of plasma of the HIV variants stand undetectable by available assays. Lifetime viral suppression improves immune activity and overall quality of life, depresses the risk of AIDS problems, and outspreads lives. ARV therapy is recommended to be utilized by everyone infected with HIV. ARVs aid folks diseased with HIV live longer, healthier and joyous lives and lessens the danger of HIV spread .

The initial advanced operative drug in the management of HIV was the NRTI, zidovudine which was sanctioned by the US FDA. Afterwards, other numerous nucleosides were advanced and three of them were pooled together to maximally suppress the virus (World Health Organization and UNITAID 2017). However, after utilization of these combined drugs amongst the infected individuals, their victory is intimidated by the advancement of viral mutations leading to HIV-1 drug-resistance. Combined drug regimens are referred to highly active antiretroviral therapy (HAART) which include 3-drug therapy. The HAART was quickly integrated into clinical utilization and swiftly showed inspiring benefit with a 60% to 80% decline in rates of AIDS, demise, and hospitalization (World Health Organization 2006, 2013a). This has remained the regular treatment option of HIV/AIDS globally.

The HAART treatment considered in Kenya include TDF_3TC_EFV/NVP or AZT_3TC_NVP/EFV as recommended first-line antiretroviral regimes combinations in treatment of HIV among ART-naïve grownups and adolescents while AZT_3TC_LPV/r or ATV/r, TDF_3TC_LPV/r or ATV/r and TDF_3TC_LPV/r or ATV/r are commended as second-line antiretroviral regimen combination (World Health Organization 2013a). Fixed dose ARV drug regimen combinations are used to reduce the pill load and boost optimal adherence (National AIDS and STI Control Program 2012). The national response team has established programs for the delivery of ART in health facilities in all the Counties of Kenya. Nonetheless, since ART was scaled up in Kenya and especially in Busia whose HIV is endemic and ART coverage stands at 95% more than the national (75%), it was important to execute this study to disclose the current situation on HIV-1 drug-resistance.

2.6 Mutation Associated with HIV-1 Drug Resistance

Mutation associated with HIV-1 drug resistance are mutations which encrypt resistance against available ARVs or classes of ARVs. They are grouped into NRTIs, NNRTIs, PIs and integrase inhibitors (II). In a study executed among ART inexperienced persons in 40 United States Cities, 0% of the individuals had major mutations encoding resistance to NRTIs, NNRTIs and PIs. However, 7% and 4% had minor mutations against NNRTIs and NNRTIs respectively (Lisa *et al.*, 2015). In a study implemented in Cameroon, 30%, 34.2% and 10.1% of ART-experienced participants had NRTI, NNRTI and PI associated mutations respectively (Meriki *et al.*, 2019).

Example of NRTIs associated mutations include: K70E/G/Q/T/N/S, M41L, D67N, K70R, L210W, V118I, T215Y/F, V75T/M/A/S, K219Q/E, M184V/I, K65R, D67G/E, K219N/R, E40F, E44D/A, L74V/I, Y115F, Q151M, N348I (Stanford University HIV drug resistance

database 2021b). Studies conducted among children and adults in Sub-Saharan Africa identified M184V as the commonest mutation, others included K65R and Q151M (Sigaloff *et al.,* 2013). Another Study in Australia identified K65R, M184V/I, T215Y/F, T215rev, M41L, L210W NRTIs associated mutations (Nicolas *et al.,* 2017). Another study in Cameroon identified M184V and K103N (Meriki *et al.,* 2019). A further study in Tanzania executed amongst ART experienced adults identified NRTIs associated mutations M184V, D67N, T69D, T215YF, K219QE, K70R, D67N, M41L, and L210W (Shabani *et al.,* 2020). Another study conducted in Uganda identified M184V in 20.7%, K65R in 8.0%, M41L in 8.0%, and K70R in 8.0% of the participant (Ivan *et al.,* 2016). NRTI associated mutations included M184V, M41L, D67N, K219Q, T215F, K65R among others (Hammers *et al.,* 2012, Kinyua *et al.,* 2017, Osman *et al.,* 2013) among HIV-1 infected folks on ART.

Example of NNRTIs associated mutations include E138K/A/G/Q/R, K103N/S/H/ T/R/Q/E, V106A/M/I, K101E/H/P/Q/R/N, V108I, V90I, L100I,A98G, I132M/L, P236L, V179D/E/F/I/L/T, Y181C/I/V, Y318F,Y188L/C/H, G190A/S/E/Q, H221Y, P225H, F227L/C/I/V, M230L/I, Y232H, L234I, K238T/N, N348I (Stanford University HIV drug resistance database 2021). Studies conducted amongst children and adults in Sub-Saharan Africa identified K103N mutation in 37% of the participants. Other commonly documented mutations were Y181C/I in16%, G190A/S in 13% and G190A in 13%, V106A/M in 10% participants (Sigaloff *et al.,* 2013). Additional study in Cameroon identified 11 NNRTI-associated mutations, T215FY, K70R D67N, L210W, M41L, and K219Q/E (Meriki *et al.,* 2019). A further cohort study in Tanzania identified K103N, V106M and G190A NNRTIs associated mutations (Shabani *et al.,* 2020). Another study conducted in Uganda identified G190A in 7.0%, K103N in 19.0%, and Y181C in 6.0% and a non-polymorphic accessory-

mutation A98G in 12.0% of the participants (Ivan *et al.*, 2016). Among the major NNRTIs associated mutations identified in Kenya included Y181C, K103N, G190A, Y188L among ART experienced individuals (Hammers *et al.*, 2012, Kinyua *et al.*, 2018, Osman *et al.*, 2013).

Example of PIs associated mutations include L90M, L89V/T, N88D/S/T/G, I85V, L76V, T74P, G73S/T/C/A/D/V, A71V/T/I/L, Q58E, I54V/A/S/T/L/M, F53L/Y, I50V/L, I50V/L, I47V/A, M46I/L/V, M36I, L33F, V32I, D30N, L24I/F/M, L23I, K20R/I/M/T/V, V11I/L, L10F/I/V/R/Y (Stanford University HIV drug resistance database 2021c). Studies conducted in among children and adults in Sub-Saharan Africa identified V82A in 17%, L90M in 15%, and I54V in 12% (Sigaloff *et al.*, 2013). A further study in Cameroon identified I54T-V82A, M46I and M46LV major PIs associated mutations (Meriki *et al.*, 2019). One participant who had been taking LPV/r in Tanzania, harbored the virus carrying L33F, L24I, V82A, I54V, N88S major PI associated mutations (Shabani *et al.*, 2020). Another implemented in Uganda identified V82A in 7.0%, M46I, I54V, and L33I in 5.0% and common accessory PI-mutations L10I in 27%, L10V in 12.0% and L10F in 5.0% were identified as well (Ivan *et al.*, 2016). In Kenya, identified major PI associated mutation include D30N, M46I, V82F (Sampathkumar *et al.*, 2013), L90M, M46I, D30N (Budambula *et al.*, 2017).

Other major Integrase associated mutations include H51Y, T66A/I/K, L74M/I/F, 92Q/G/V, T97A, G118R, G118R, E138K/A/T, G140S/A/C P142T, Y143C/R/H/K/ S/G/A, P145S, P145S, S147G, Q148H/K/R/N, G149A, V151I/L/A, S153Y/F, N155H/S/T/D E157Q, G163R/K, S230R, R263K (Stanford University HIV drug resistance database 2021a). In a study conducted in India 0% of participants had Major Integrase associated mutations,

however, 13.8% had E157Q polymorphic mutation and 1.7% had L74IM, Q95K, T97A accessory mutations (Santosh *et al.,* 2019). Studies conducted in Kenya have not exhaustively identified all the mutations encoding resistance against all the four classes of ARVs, therefore the current study was planned to identify other new mutations that are associated with HIV-1 drug resistance which were not previously identified in the preceding studies.

CHAPTER THREE

RESEARCH METHODS

3.1 Study Area

The present research was implemented in the County of Busia (Figure 2). Three health amenities (Busia County health facility, Matayos health Centre; and Khunyangu Sub County health facility) were considered in the current study. The latitude of the county is $0^{o}27'36.18''$ towards the North of the Equator and the longitude is $34^{o}06'42.08''$ towards the East of the equator. The County lies at the borders of Kenya-Uganda towards the East and North of Kakamega and Bungoma Counties respectively. It is also located at the South of Siaya County and Lake Victoria. Neighbors Uganda towards the West. The county relies heavily on agriculture and fishing as commercial activities. The county has a population size of 893,681 (Kenya National Bureau of Statistics 2019) both male and female across all ages. The County is amongst top five with a high HIV and AIDS prevalence of 7.7%, 38606, after Migori, 13.3%; Kisumu, 16.3 %; Homabay, 20.7%; Siaya, 21.7 % in Appendix 12 (Ministry of Health and National AIDS Control Council 2018b). Of the HIV-1 diseased individuals in the County of Busia, 95%, 33654 are utilizing antiretroviral. The coverage of utilizing ART coverage is greater in the County of Busia comparable with 61%; 62 %; 76% and 71% correspondingly for the four Counties with uppermost prevalence of HIV (National AIDs Control Council 2018a, b, c, Onyango *et al.*, 2019)

The County of Busia Map

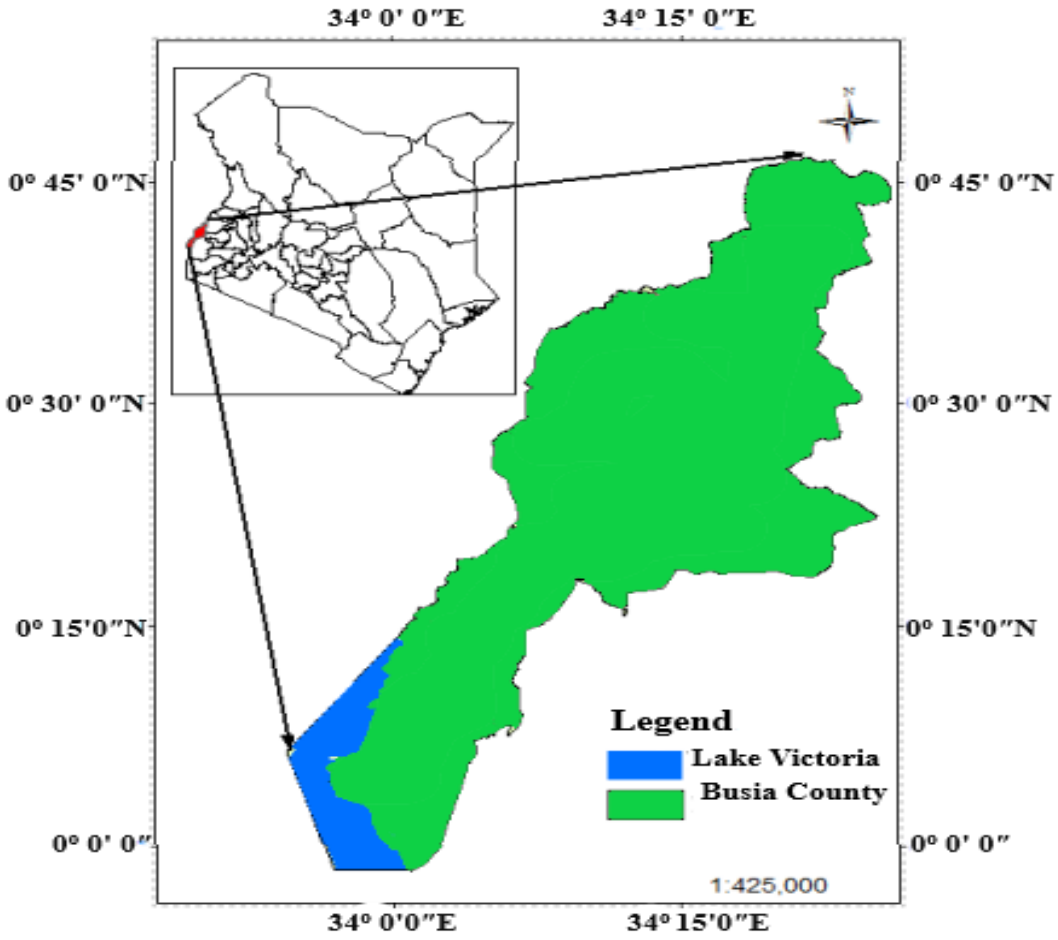


Figure 2: - Study area. The of Busia map where the study was implemented. Source: http:// www. jkuat. ac. ke/ departments /geg is /wp-content /uploads / 2016/ 05/ N1004 01104112.pdf

3.2 The Study Design

The study employed cross sectional study design of which blood samples were gathered during the months of July, 2019 – June, 2020 for analysis of HIV-1 viral tallies of the participants during the third visit at the health facilities, subtypes and drug resistance. Retrospectively, info on the history of the preceding kind of ARVs utilization and Viral tallies of the participants for the first two sequential visits before the start of laboratory

procedures was gathered from the NASCOP catalogued database during the months of January- June, 2019. This NASCOP database is connected to KEMRI domain. KEMRI is the lone producer of all data on the viral loads data on behalf of NASCOP. Advantages of cross sectional studies include time saving and cost effective comparable to case control studies. Supplementing on this, info on the history of ARVs uptake was collected retrospectively to minimize biases and evaluate the risk of a population.

3.3 Research Population

This involved infected participants with HIV who were utilizing or/and not utilizing ARVs in the County of Busia.

3.3.0 Inclusion-Criteria

Participants of ≥ 18 years of age, resided in the County of Busia, who had utilized ARVs for at-least 6 months' period or/and not utilizing ARVs and consented to take part in the study were included in the study.

3.3.1 Exclusion-Criteria

Individuals of < 18 years of age, taken ARVs for less than 6 months or/and not utilizing ARVs and resided in the County of Busia for less than five years and not willing to consent to be part of the study.

3.4 Sample Size Calculation

Formulae by (Kothari 2004) was utilized to compute minimum number of sample size

$$n = \frac{N}{1+N\,(0.05)^2}$$ Where

$$n = \frac{893{,}681}{1+893{,}681(0.05)^2}$$

N stands for a population size of Busia County which is 893,681 by the year 2018 (Kenya National Bureau of Statistics 2019); n = minimum sample size; 0.05 represents margin error. By substituting these parameters in the above formulae, this counts for 400 samples which represent the whole population (HIV-1 positive and HIV-1 negative individuals) in Busia County. However, since the study included only HIV-1 positive cases as per the inclusion criteria, the above samples were halved into two to give 200 blood samples. Out of the 200 blood samples, 95% which is ART coverage for the county =190 represents those taking ARVs and 5% =10 those not taking ARVs.

3.5 Sampling-Technique

Probability sampling-technique was employed for the current study in a manner that, participants were engaged systematically.

3.6 Sample-Collection

The sample-collection utilized the following practice (Jordon *et al.,* 2012). Retrospective data on the ART utilization was gathered during the months of January – June, 2019. Cross section ally, samples were gathered during the months of July 2019 to June 2020. It was vital to appreciate the kind of ARVs participants were utilizing before the study, to inaugurate whether swapping to other ARV treatment had occurred or not. Data about the duration of previous ART utilization, gender, and age were gathered on a questionnaire shown in appendix 11. About, 5 milliliter of sample of blood were drawn using sterilized syringes and needles. This was poured in EDTA tubes. Triple packaging of the blood samples was executed and transported to the laboratory of CIPDCR, KEMRI following the standard nationwide sample conveyance program for clinical samples. Blood samples were detached from the cold environment and allowed to reach at room-temperatures for

approximately one hr. before analyzing. The quality of blood samples was inspected for hemolysis, clotted blood was excluded. The correct amount, 5 milliliters was also inspected. Plasma was separated from whole-blood and stored -80ºC. Quantification of viral counts was executed at CIPDCR, KEMRI. Plasma for extraction of RNA and sequencing were transported to KEMRI Nairobi for investigation.

3.7 Sample Processing

3.7.1 RNA Extraction/Viral Load Quantification using nested-PCR

Extraction of RNA is the process of which the RNA material of HIV-1 is made available by lysing of the cell and nucleus HIV. Extraction of RNA was carried out and quantified with the aid of Abbott *m2000* sample preparation grounding method (Abbott Molecular and Des Plaines 2014). About, 0.2 milliliter of HIV-1 assay extraction mixtures was employed according to manufacturer's regulations (Abbott Molecular and Des Plaines 2014). Counting of viral tallies was executed by aid of real time PCR. Plasma samples were thawed at ambient temperature. Plasma samples were obtained through centrifuging complete blood in green colored toped EDTA tubes. About 400 micro liter of samples were re-conveyed into tubes of PCR. Lysis of the plasma samples executed by aid of 10% tween buffer and 4.7m guanidium iso-thiocyanate. Positively charged micro particles played a role of attracting the RNA. Elimination of impurities from captured RNA-complex was done using buffers, wash 1 consisting of GITC/tween and wash 2 comprising of RNAse free -water were utilized. Elution buffer containing great salt concentration was utilized to separate the RNA with micro-particles and displacing the RNA-molecules in PCR tubes.

3.7.2 Reverse Transcription by PCR

Reverse-transcription by PCR is the assay by which HIV-1 complementary deoxyribonucleic nucleic acid (cDNA) is transformed from RNA for further analysis. The transformed cDNA is steady in outside environments comparable to RNA. The transcription was carried out with the aid of commercially acquired real time PCR procedure by Thermo-Fisher's Scientific's-Genotyping-kit (ThermoFishers Scientific 2017)**.** Initially, the Real Time PCR reaction mixtures consisting of the Real time PCR master mixtures, normal and Super-Script iii enzyme was fixed for the requisite amount of reactions. The obtained RNA, positive (a sample consisting of viral tallies of 1000cp/ml and more) and negative controls were denatured using a thermocycler (Veriti™-96-Well-Thermal -cycler) at 65°C for 10 minutes. This process changed the minor bonds within the RNA making single strand-RNA. About, 10 micro-liter of the obtained RNA was accompanied with 40 µl of the master-mixture involving primers of RT-PCR, Super-script III, MgSO4, dNTPs and Platinum-Taq, RT-PCR mix (1 micro-liter of the enzyme of Super-Script III, 39 micro-liter of the Mix of RT PCR. About 50 micro-liter comprehensive volume was reverse-transcribed. The enzyme of superscription speeds up the reverse-transcription of cDNA from RNA.

The cDNA was employed as an outline for PCR. Catalysis using Platinum Taq was employed for the cDNA amplification. The PCR reverse transcription was executed under the subsequent cycling environments as per manufactures regulations: one round of reverse transcription cycle at 50^0C for 45 minutes, one round of enzyme inactivation cycle at 94^0C for 2 minutes, 40 rounds of denaturation cycles at 94^0C for 15 seconds. 40 rounds of annealing cycles for at 50^0C for 20 seconds, 40 rounds of extension cycles at 72^0C for 2

minutes and one round of final extension cycle at 72^0C for 10 minutes. The extracted RNA were finally computed. About 1.8Kb was the end PCR product (quantified RNA – Viral load)

3.7.3 Amplifying cDNA by aid of Nested-PCR

The Thermo-Fisher Scientific's-Genotyping kit was acquired commercially (ThermoFishers Scientific 2017) for use during amplification of cDNA. The reverse-transcribed PCR 1.1kb cDNA was amplified under the nested PCR. Sufficient nested-PCR-reaction mixtures (Amplification Thermus aquaticus Gold-LD DNA polymerase-enzyme and Nested-PCR mixture) was primed for the essential amount of reactions. Approximately 2 micro-liter of the real time PCR products was supplemented to 48 micro liters of the nested PCR master mixtures (About 47.5 micro liter of the Nested PCR mixtures (PCR-primers; MgCl2; dNTPs; PCR buffers), and about 0.5 µl Amplification Thermus aquaticus Gold-LD DNA polymerase-enzyme). An overall volume 50 micro-liter was utilized. The nested-PCR was executed under the preceding rounding conditions following the manufactures' rules: one round of initial denaturation cycle at 94^0C for 4 min, 40 rounds of denaturation cycles at 94^0C for 15 secs, 40 rounds of annealing cycles at 55^0C for 20 secs, 40 rounds of extension cycles at 72^0C for 2 min, one round of final extension cycle for at 72^0C 10 min. Amplification Thermus aquaticus Gold-LD-DNA polymerase enzyme catalyzed PCR product amplification. The final Nested-PCR-product generated was 1.08 kb.

3.7.4 Confirmation of PCR Product by Gel-Electrophoresis

Gel electrophoresis procedure was employed to determine the presence of cleaned Nested PCR product of amplicon. The nested PCR amplicon was 1.1 kb. Approximately 1% gel of agarose was primed following these methods, about 1g of agarose and 100 milliliter of IX-TAE buffer were mixed in a conical bottomed flask. The content was warmed for

approximately 80 secs at temperatures between 50ºC - 60ºC, the mixture was further whirled and cooled to a room temperature condition. About 5 micro liter of ethidium-bromide was conveyed into 100 milliliter of agarose mixes, this were further mixed exhaustively. The trays of the gel and combs were set and the gel of agarose was transferred prudently into the caster of gel. Approximately 5 micro liter of 6X-DNA packing buffer was complemented onto 5 micro liter of the sample onto a Para film and mixed exhaustively. In each and every well 10 micro liter of the reaction mixes and 5 micro liter of bromophenol blue was complemented, in the initial well, 4 micro liter of a DNA ladder was supplemented. The contents in the gel was allowed to run at 100V, then further soaked in buffer comprising ethidium bromide for one hour. The amplicons in gel were photo'd at 320 Nano-meter and imagined by use of the imaging system and photo'd (Visi-Doc-It™ Imaging-System, CA/USA) (Figure 3). The cDNA confirmed product was 1.1kb in size.

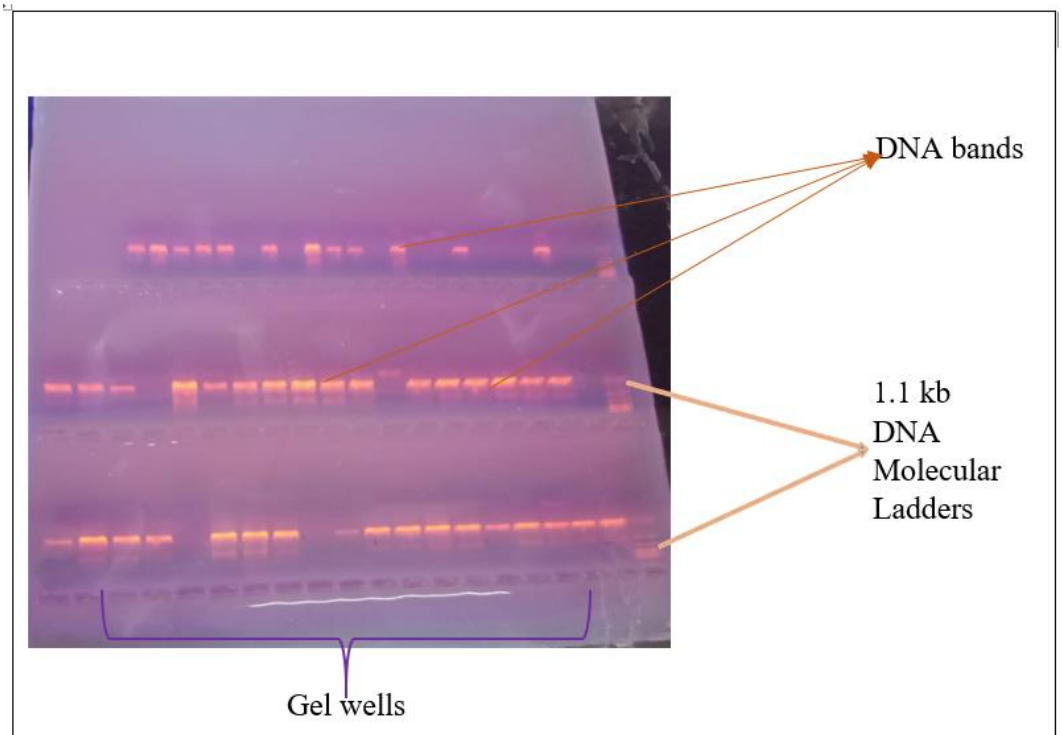


Figure 3: Photograph of confirmed cDNA in Gel visualized under Imaging System

The entire samples were duplicated, those with unsuccessful bands at the validation level on 1% of agarose-gel, the processes were continual just away from the RNA-extraction to confirmation.

3.7.5 Purifying of Nested cDNA PCR product

Cleaning of the cDNA PCR product was accomplished by aid of commercially acquired Thermo Fisher-Scientific's Clean-Sweep Purification Component basing on manufacturers' instructions (ThermoFishers Scientific 2017). About 6.4 micro liter of PCR ExoSAPIT™ Product Cleanup reagent was added to 14 micro liters of each PCR Nested product. The contents were vortexed for about two - three sec, and then centrifuged at 12000 revolutions per minute. Denaturation of cDNA fragments of ≥300bp was done using the Clean Sweep reagents. Approximately, 20.4 micro liter reaction mixes were cleansed under the succeeding conditions: Digest at 37^0C for 15 min and heat deactivation at 80^0C for 15 min.

3.7.6 Cycle-Sequencing

Cycle-sequencing was executed using a total of sequence of 6 primers (Forward and Reverse 1; 2 and 3). These were sourced commercially (ThermoFishers Scientific 2017). About 3 micro liter of the cleansed PCR nested products were supplemented into about 18 micro liter of each and every six primer mixes (sequencing buffer, ddNTPS, Thermus aquaticus polymerase enzyme, primer, dNTPs). Each and every primer was poured into a separate well onto a plate of PCR. To each and every well, a small sample mixes of the purified PCR product was added. The targeted regions (protease, Reverse transcriptase, and non-reverse transcriptase) of the pol region of HIV-1 were amplified with the listed primers. About, 21 micro liter of the positive sequencing control (sample with HIV-1

mutations at the reverse and protease region of the pol gene) and PGEM (Appendix 4) which consisted cDNA, primers and terminator were further added to at least one well. The cycle-sequencing environments were fixed following the manufactures guidelines: 25 rounds of denaturation cycles, annealing cycles, and extension cycles at 96°C for 10secs, at 50˚C for 5 seconds, at 60˚C for 4 minutes respectively. The pol genetic factor of HIV type1 was amplified and investigated for mutations and subtypes since it is highly preserved and carry the major three enzymes (protease, reverse transcriptase and integrase) liable for reproduction succession of the HIV virus. Furthermore, all the accessible and available antiretroviral drugs mark these three major enzymes for suppressing of the virus.

3.7.7 Purification Cycle Sequenced products

Purifying of Cycle sequenced products was executed with the aid of Thermo Fisher Scientific Big Dye X-Terminator purifying Kit (ThermoFishers Scientific 2017). The X-terminator mixes stored at 4˚C was detached from the cold environments and set to equalize to ambient temperatures. These were vortexed at approximately 10 secs. About 20 micro liter of an operational solution of Big Dye XTerminator mixes and 90 micro liter of the SAM mixes solution was primed to clean 20 micro liter of the cycle sequence products. Approximately, 110 micro liter of SAM-Big Dye the X-terminator mixes were conveyed in each sample in vessels of PCR and vigorously mixed well. The Sequencer plate was enveloped by aid of micro Amplification sterile adhesive-film. The plates with mixes were vortexed for 30 minutes at 7000rpm and centrifuged at ambient temperature for 1000rpm. Cleansing of the sequencing products were executed to eradicate un-incorporated ddNTPs. Kit's optimization had previously been executed by Thermo-Fisher for Sanger sequencer platform on HIV-1 (ThermoFishers Scientific 2017).

3.7.8 Sequencing and Sequence Analysis

The ABI3500 genetic-analyzer was employed for sequencing procedure. Acquisition of the sequence scanner software was from www. Thermo-fisher. Com and used to give verdicts of quality of the sequence. Sequencing was executed by Sanger sequencer from Massachusetts and Waltham, USA following this procedure: The 96 well plated septa was concealed with adhesive film before conveying onto the 3730 XL-DNA analyzer for sequencing directly. Dye Set Z Big Dye V3 and Run module long Seq 50 POP 7/1 were employed.

Sequencing procedure was executed to describe the nucleoside bases sequence in an amplified output. Fluorescently trademarked dyes were linked to ACGT elongation products in the complementary DNA sequencing reactions utilizing the Applied-Bio systems Sequencing kits 3.1 version (Thermo Fisher Scientific 2018). The dyes were trademarked to brand the bases as black, Guanine; blue, Cytosine; green, Adenine; red, Thymidine. The end 5'dye-label primers and 3'dye-labeled di-deoxy-nucleotide terminators were used to fuse the dyes. Chief technical restrictions like sequenced bases' quality electro-pherograms and the raw-data, were appraised by Seq Scanner version 6 (Applied-Bio systems Inc. USA). Each and every amplified segmented sequence generated by the Sanger-sequencer were assembled, aligned to yield consensus-sequences (Appendix-5) by aid of Recall (beta version 3.05), a software (http://pssm.cfenet.ubc.ca/) manufactured by the excellence center for HIV research at the British University of Columbia. The software further generated DNA sequences in Fasta files (Appendix-6). The software again produced drug-resistance report (Appendix-7) of each and every sample.

Each and every sequences of Fasta format generated by the Recall-software were further deposited into the Database of Stanford University HIV Drug Resistance (http:// sierra2. stanford. edu/ sierra/servlet / JSierra) which analyzed and interpreted mutations encoding resistance to specific ARVs (Appendix 8). This database standard and accurate tool for identifying HIV-RT and HIV-PI sequences and HIV drug resistant mutations.

3.7.9 HIV type 1 subtypes analysis

Sequences in Fasta form of each sample generated by Recall software were singly dropped into the jumping-profile Hidden-Markov Model (http://jphmm.gobics.de) software which further inaugurated HIV-1 subtypes and other recombinant (Schultz *et al.,* 2009, Zhang *et al.,* 2006) (Appendix-9). Sample sequence of one HIV-1 reference DNA was obtained from the database NCBI (https: // blast. ncbi. nlm. nih. gov) and one Hepatitis 'C', an out- group DNA sequence was also obtained from NCBI database for use in the analysis of phylogenetic relationship (https: // blast. ncbi. nlm. nih. gov). Both Pairwise/multiple alignment of the two reference sequences and the sample sequences were executed by help of ClustalW (Appendix-10). MEGA7 maximum-likelihood process was utilized to build a phylogenetic-tree (Kumar, Stecher and Tamura, 2016) (Figure 6). KU749431.1.14……HIV-1 and Hepatitis C Nucleotide sequence was retrieved from NCBI

3.8 Data Analysis

Congregated data was entered and stored excel Microsoft sheet. The analysis was executed by help of SPSS ver. 20 (Armonk NY: IBM-Corp). Descriptive tests were employed to analyze socio-demographic factors and the frequency of different mutations. Chi square was utilized to compute the relations of two measurable variables. Measureable variables encompassed the health facility, gender, ART-uptake duration, viral tallies. Systematic

reviews were executed where by 100 related manuscripts were obtained through MED-LINE and 23 of them were eligible to be included in the study

3.9 Ethical Clearance and Considerations

Approvals for the current study were obtained from the Ethical and Scientific authorities for national ethical-review committee where by the proposal was assigned a study number (KEMRI_SERU_CIPDCR_008_3333) by the Kenya Medical Research Institute. The approval for this study was similarly obtained from the Department of postgraduate studies, DPS at Masinde Muliro University of Science and Technology (appendixes 1, 2, 3). The study including its importance was explained to all the participants. The participants signed the cognizant consent. Skilled persons at the health facilities collected the samples. The identifiers of the participants were maintained secret instead the sample code was utilized to brand the sample for confidentiality purposes. All the outcomes on drug resistance and sub-types were shared with the clinical personnel directly linked to participants in all the facilities so that to utilize the results in the management of participants appropriately. If there was an existence of mutations in a sample of participants that encrypted resistance to two or three drug regimen-combination, the clinicians opted to change the participants to an alternative appropriate regimen.

CHAPTER FOUR

RESULTS

4.1 Sociodemographic and Clinical Characteristics

Minimally, two hundred participants were enrolled for the study. Out of these, a total of 160 samples obtained from corresponding number of participants conceded sequencing process. Female on ART and not on ART were 92, 57.5% and 6, 3.8% respectively. Male on ART and not on ART were 58, 36.2% and 4, 2.5% respectively. The general mean age of participants was 32.7 years with the youngest and the oldest being 18 and 70 years respectively. The overall average viral loads were 88052 (min. 668 and max. 2623390) copies of viruses per milliliter of blood. General mean-duration of ARVs usage was 6.34 (minimum, 0.5 and maximum, 14) years (Table 1).

Generally, an overall of nine different drug-regimen combinations (5 and 4, first and regimen respectively) were being consumed by all the participants. There was arithmetical importance difference (χ 2=29.463, p=0.021) in the utilization of various types of drug-regimen combination amongst hospitals in that participants at Matayos were utilizing entirely nine drug-regimen combinations comparable to participants in Khunyangu and Busia who were utilizing a total of 7 each. An intention which warrant more exploration. Most of the participants were utilizing the reverse transcriptase and protease inhibitors based pooled drug-regimen 138, 92% compared to integrase pooled drug-regimen 12, 8% simply because integrase pooled drug regimen was presented in Kenya lately (World Health organization and UNITAID 2017) (Table 2).

Table 1: Socio-Demographic and Clinical Characteristics

	Health facilities			
	General n=160	Busia n=54	Khunyangu n=53	Matayos n=53
Female n(%) on ART	92(57.5)	31(54.4)	34(64.2)	27(50.9)
Female not on ART	6(3.8)	2(3.7)	2 (3.8)	2(3.8)
Male n(%) on ART	58(36.2)	19(35.9)	16(30.2)	23(43.4)
Male not on ART	4(2.5)	2(3.7)	1(1.9)	1(1.9)
Mean-age(min-max) years	32.7 (18-70)	35.8 (18-66)	29.3 (18-59)	33.12 (18-70)
Mean viral loads (min-max) copies/mm^3	88052 (668-2623390)	56065 (668-651,800)	86995 (1878-714,628)	118,097 (1118-2,623,390)
Mean duration on ART uptake (min-max) years	6.34 (0.5-14)	5.89 (0.5-14)	6.24 (1-14)	6.9 (1-14)

Table legend: The report here includes information of 150 HIV-1 infected participants on ART for a duration of 6 months and above and 10 HIV-1 infected and not on ART whose samples proceeded to sequencing level. Women were somewhat more comparable to men. The youngest and the oldest participants were 18 and 70 years old. Viral tallies of the participants were computed by help of Abbott m2000 sample preparation platform. Descriptive test was employed to compute the frequencies of the above variables

Table 2: Uptake of ARVs Drug-Regimen Combinations

Drug-regimen combinations	Health facilities				χ^2, p value
	Over-all, n=150,%	Busia, n=50,%	Khunyangu, n=50,%	Matayos, n=50,%	
AZT;3TC;ATV/r	20,13.3	8, 16.0	7, 14.0	5,10.0	
AZT;3TC;EFV	21,14.0	0, 0.0	8, 16.0	13, 26.0	
AZT;3TC;LPV/r	7,4.7	0, 0.0	5, 10.0	2, 4.0	
AZT;3TC;NVP	35, 23.3	11, 22.0	13, 26.0	11, 22.0	
TDF;3TC;ATV/r	13, 8.6	5, 10.0	4, 8.0	4, 8.0	29.463, 0.021
TDF;3TC;DTG	12, 8.0	3, 6.0	4, 8.0	5, 10.0	
TDF;3TC;EFV	36, 24.0	19, 38.0	9, 18.0	8, 16.0	
TDF;3TC;LPV/r	3, 2.0	2, 4.0	0, 0.0	1, 2.0	
TDF;3TC;NVP	3, 2.0	2, 4.0	0, 0.0	1, 2.0	

Table legend: The table reports information of 150 participants who were utilizing ARVs. DTG, Dolutegravir; LPV/r, Lopinavir/ritonavir; ATV/r, Atazanavir/ritonavir; 3TC, Lamivudine; AZT, Azidovudine; EFV, Efevirenz; TDF, Tenofovir-disoproxil-fumarate; NVP, Nevirapine. NVP; EFV are grouped as NNRTIs; TDF, AZT, 3TC NRTIs; LPV/r and ATV/r PIs and DTG Integrase inhibitor-drug. Pooling of three of these drugs in one drug regimen combination include: TDF;3TC;DTG, TDF;3TC;NVP/EFV, AZT;3TC;NVP, AZT;3TC;EFV are classified as 1st-line drug-regimen combinations and TDF;3TC;LPV/r AZT;3TC;ATV/r, TDF;3TC;ATV/r, AZT;3TC;LPV/r are 2nd line. Descriptive tests of SPSS v.20 was employed to calculate prevalence and frequencies throughout entire drug-regimen combinations. Chi-square, χ^2 was employed to calculate arithmetical significance of the drug-regimen combinations amongst health facilities

4.2 HIV-1 Subtypes Prevalence

Over ally, 12 dissimilar HIV-1 subtypes were identified. Sub-type A_1 99, 61.9% was dominant; followed by D with 31, 19.4%; then A_1_B plus A_1_D with 7, 4.4% each; A_1_C 6, 3.8%; A_1_G 3, 1.9%; C 2, 1.3% in this manner. Sub-types A_1_A_2, A_1_J, A_2_D, A_1_K, B_C were the least with 1, 0.6% each (Table 3). Sub-type A_1 has been reported in many studies being the dominant in Kenya equally in the current study. Earlier introduction of this sub-type from South Africa into Kenya and its speedily transmission comparable to others (Khamadi *et al.*, 2005) could explain this.

Table 3: HIV type 1 Sub-types Prevalence

Sub-type	General, n = 160(%)	Busia n = 54(%)	Khunyangu, n = 53(%)	Matayos, n = 53(%)
A_1	99(61.9)	30(55.6)	35(66.0)	34(64.2)
A_1_A_2	1(0.6)	1(1.9)	0(0.0)	0(0.0)
A_1_B	7(4.4)	4(7.4)	1(1.9)	2(3.8)
A_1_C	6(3.8)	1(1.9)	3(5.7)	2(3.8)
A_1_D	7(4.4)	1(1.9)	2(3.8)	4(7.5)
A_1_G	3(1.9)	0(0.0)	2(3.8)	1(1.9)
A_1_J	1(0.6)	1(1.9)	0(0.0)	0(0.0)
A_1_K	1(0.6)	1(1.9)	0(0.0)	0(0.0)
A_2_D	1(0.6)	1(1.9)	0(0.0)	0(0.0)
B_C	1(0.6)	0(0.0)	0(0.0)	1(1.9)
C	2(1.3)	1(1.9)	1(1.9)	0(0.0)
D	31(19.4)	13(24.1)	9 (17.0)	9(17.0)

Table legend: The HIV-1 subtypes of 160 are reported. Sequences of DNA in Fasta form listed as an appendix-6 produced by Recall (beta version 3.05) software were dropped in Jumping-Profile-Hidden Markov Model database to blast the sub-types (Appendix-9). Descriptive tests in SPSS-Version 20 was utilized to analyze the counts and the prevalence throughout the entire HIV type 1 sub-types.

The illustration below shows an example of output of Jumping Profile Hidden Markov Model database. Where a cDNA sequence in a fasta format of each and every identified HIV-1 from a sample' participant was analyzed for subtype/recombinants. The figure shows that the individual whose sample was analyzed was living with HIV-1 recombinant A_1D (Figure 4)

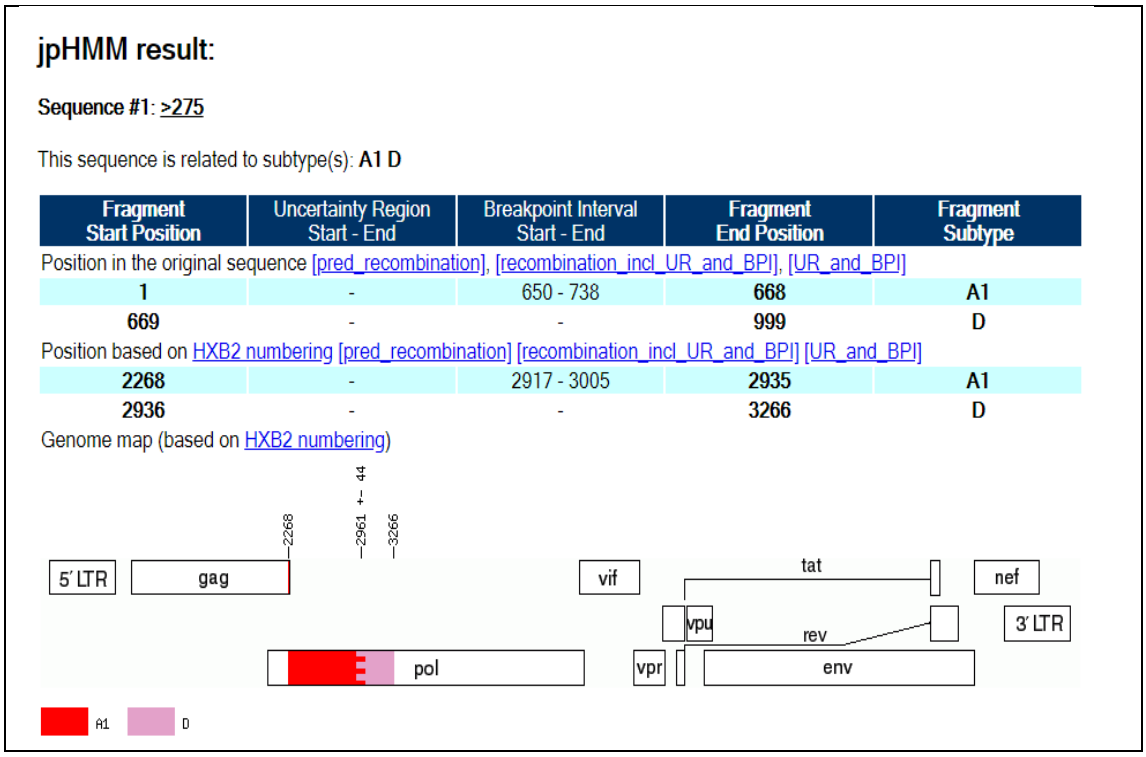


Figure 4: HIV-1 Subtype/Recombinant by Jumping Profile Hidden Markov Model database, showing HIV-1 recombinant (A_1D)

The outcomes of every cDNA sequences of nucleoside were aligned by the help of CrustalW within MEGA7 (Figure 5 and Appendix 10) database and tree of phylogeny (Figure 6) showing the relationship of samples was illustrated by aid of maximum likely-hood. There existed insertion, substitution, and deletion of some nucleotide bases in some sequences of samples comparable to the sequence of standard HIV-1, KU749431.1.1435.

Figure: 5: Alignment of nucleotide bases by CrustalW

Tree of phylogeny showing the relationship of HIV-1 subtypes/recombinants identified from participants' samples were illustrated by aid of maximum likely-hood (Figure 6) There existed a relationship between the sequences of the samples and the sequence of the standard HIV-1, KU749431.1.1435 in that they clustered together. There was no relationship of the sequences of samples and an outgroup in that sample' outcomes clustered away from an outgroup (Hepatitis C)

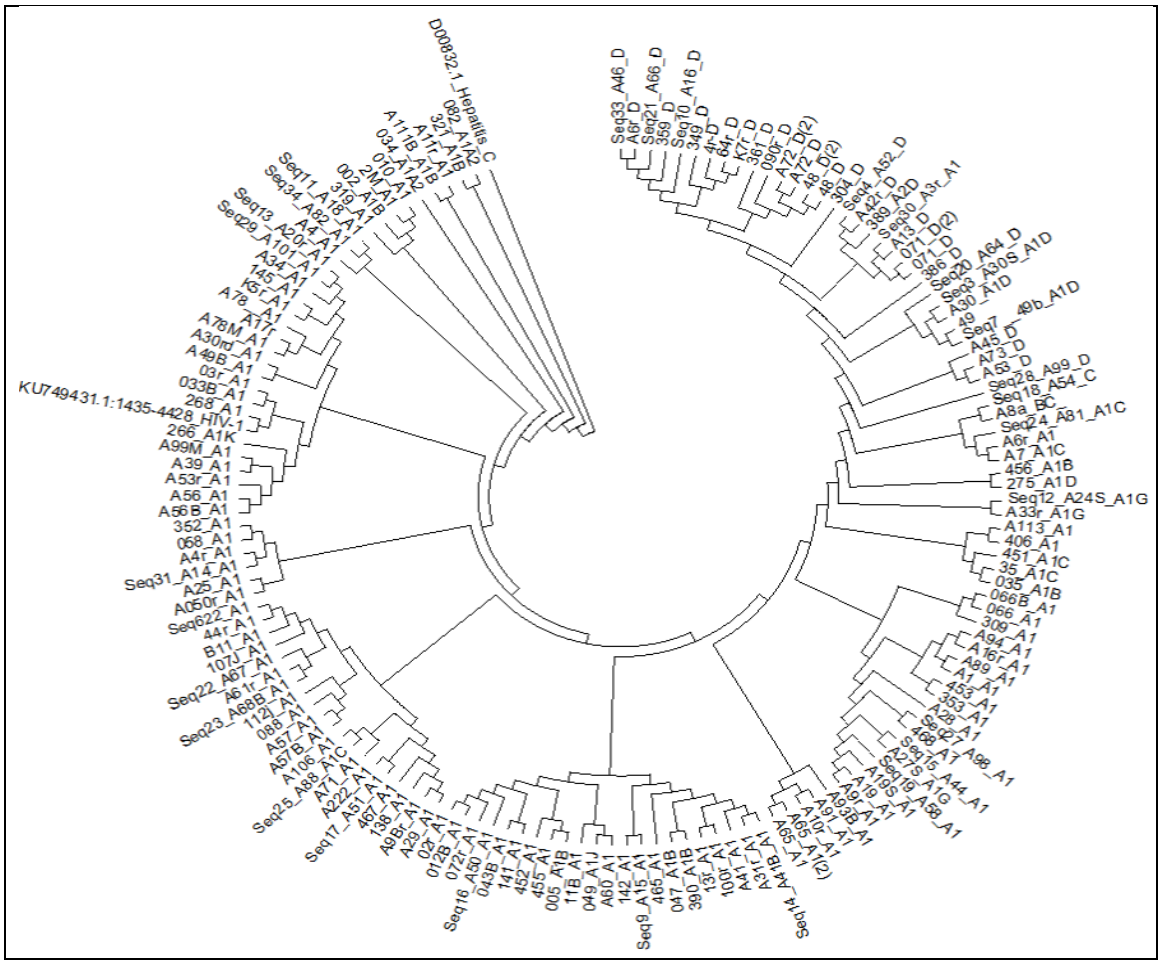


Figure 6:- Molecular Phylogenetic tree showing relationship of HIV-1 subtypes
Evolutionary history was concluded with aid of maximum-likelihood procedure grounded on the Tamura/Nei-model (Tamura and Nei 1993). Tree of the uppermost log-likelihood (-14182.95) is illustrated above. The degree of trees by which the related taxa grouped together is illustrated just next to branches. Original tree for the experimental investigation were acquired automatically by relating neighbor Joining and Bio-NJ systems to a matrix of pairwise/multiple distances computed by use of maximum composite likely-hood (MCL) approach, and then selecting the topology with superior log likely-hood significance. The investigation composed of 150 nucleotide base sequences. The codon locations encompassed 1st/2nd/3rd and noncoding. All points with < 50% coverage were rejected. The total final dataset were 999 sites. Evolutionary investigation were executed in MEGA 7 (Kumar *et al.,* 2016)

4.3 HIV-1 Drug Resistance Prevalence

In general, the prevalence of HIV-1 drug-resistance mutations amid participants was 110, 73.3%. Matayos health facility reported somewhat greater equated to other dual though there was no arithmetical implication difference. The overall prevalence of NRTIs of HIV-1 drug resistance was 88, 58.0% with Matayos again reporting more comparable to other two with arithmetical implication difference of a chi square value 6.983 and p value 0.033.

This was due to the fact that surplus participants in Matayos were utilizing AZT;$_{3TC}$;EFV/NVP of which a drug of AZT has been in utilization for more number of years comparable to TDF that is pooled into TDF;$_{3TC}$;EFV/NVP of which many participants in Busia and Khunyangu facilities were utilizing (National AIDS and STI Control Program 2011). To add on this, the frequency of mutation M184V which encode resistance to AZT predominated in this study and was higher in Matayos thus giving the facility an upper hand to dominate compared to Busia and Khunyangu (Table 6). Over ally the prevalence of HIV-1 drug-resistance of NNRTIs was 68.0% of which Busia, Matayos Khunyangu, recorded 68.0%, 78.0% and 58.0% respectively with no arithmetical implication difference, p = 0.1. HIV-1 drug-resistance of PIs was 10.7% over ally of which Busia, Matayos, Khunyangu recorded 10.0%, 8.0% 14.0% respectively with no arithmetical implication difference, p=0.716 (Table-4).

In the present study, more participants were consuming 1st line drug regimen combinations, 107 comparable to the 2nd line drug regimen combinations, 43. HIV-1 drug resistance prevalence amongst the participants taking the 1st line drug regimen combination was 70.1% and the 2nd line drug regimen combination was 8.6% (Table-4). The fact that HIV-1 infected persons are initiated on the 1st line drug regimen combinations prior to switching them on the 2nd line drug regimen combinations (National AIDS and STI Control Program 2016, World Health Organization 2018) could define this scenario.

Table 4: HIV type 1 Drug-Resistance Prevalence

	Health Facilities				
	Over-all, n=150 each, %	Busia, n=50 for each, %	Khunyangu, n=50 for each, %	Matayos, n=50 for each, %	χ^2, P value
Mutation					
Present	110, 73.3	37, 74.0	34, 68.0	39, 78.0	1.295, 0.523
NRTIs	88, 58.7	29, 58.0	23, 46.0	36, 72.0	6.983, 0.033
NNRTIs	102, 68.0	34, 68.0	29, 58.0	39, 78.0	4.596, 0.1
PIs	16, 10.7	5, 10.0	7, 14.0	4, 8.0	0.979, 0.716
	n =107, %	n = 35, %	n =34, %	n =38, %	
1st line	75, 70.1	25, 71.4	17, 50.0	29, 76.6	
	n = 43,%	n=15, %	n =16, %	n =12, %	
2nd line	8, 18.6%	2, 13.3	4, 25.0	2, 16.7	

Table legend: NRTIs_nucleotide-reverse-transcriptase-inhibitors; NNRTIs_non; PIs_protease inhibitors. NNRTIs namely, NVP and EFV forms part of 1st line antiretroviral. NRTIs namely, TDF, 3TC, AZT, and Integrase DTG forms part of both the 1st and the 2nd line antiretroviral. PIs namely, ATV/r and LPV/r forms part of 2nd- line drug regimen combination. These individual drugs are pooled into three to form HAART as follows: TDF,3TC,DTG; TDF,3TC,NVP, AZT,3TC,NVP; TDF,3TC,EFV; AZT,3TC,EFV are referred to as 1st -line drug regimen combinations AZT,3TC,ATV/r; AZT,3TC,LPV/r; TDF,3TC,ATV/r; TDF,3TC,LPV/r are referred to as 2nd -line drug regimen combinations. Descriptive tests in SPSS v. 20 was employed to calculate the prevalence throughout the entire HIV-1drug-resistant mutations and 1st and 2nd lie drug regimen combination. Chi square, χ^2 was utilized to compare the prevalence of drug resistance across the three health facilities.

4.4 Patterns Prevalence of HIV type1 Drug Resistance Mutations

HIV type 1 drug-resistance patterns prevalence was described in Table 5. Over-ally, the HIV-1 drug resistance pattern prevalence against dual-class antiretroviral were listed as follows: NRTIs/NNRTIs, 46.7%; NRTIs/PIs, 0.7%; NNRTIs/PIs, 0.7%. The pattern prevalence against multi class antiretroviral NRTIs/NNRTIs/PIs was 8.0%. The pattern prevalence against single-class antiretroviral were NNRTIs 17, 11.3%; NRTIs 7, 4.7% and PIs 2, 1.4%. Extra pattern of HIV type 1 drug-resistance prevalence was observed for dual-class NRTIs/NNRTIs comparable to other class. The fact that the two class of drugs were initiated for use in the country earlier compared to PIs could explain this. Additionally, all the HIV-1 infected persons are initiated on the first line drug regimen combination prior to changing them to second line drug regimen combinations (National AIDS and STI Control Program 2016, World Health Organization 2018).

Table 5: Patterns Prevalence of HIV type 1 Drug-Resistance

	Health-Facilities			
	Over-all, n=150 ,%	Busia, n=50,%	Khunyangu, n=50,%	Matayos, n= 50, %
NNRTIs/PIs	1, 0.7	1, 2.0	0, 0.0	0, 0.0
NRTIs/NNRTIs	70, 46.7	23, 46.0	16, 32.0	31, 62.0
NRTIs/NNRTIs/PIs	12, 8.0	3, 6.0	4, 8.0	5, 10.0
NRTIs/PIs	1, 0.7	5, 10.0	0, 0.0	1, 2.0
NNRTIs	17, 11.3	5, 10.0	9, 18.0	3, 6.0
NRTIs	7, 4.7	5, 10.0	2, 4.0	0, 0.0
PIs	2, 1.4	0, 0.0	2, 4.0	0, 0.0

Table legend: The table report the prevalence of patterns of HIV-1 subtype. This means that some participant's samples harbored viruses encoding resistance to more than one class of ARVs. NRTIs_nucleotide-reverse transcriptase-inhibitors; NNRTIs_non; PIs_protease inhibitors. Descriptive tests in SPSS v.20 was utilized to calculate the of pattern prevalence of HIV type 1 drug-resistance across the entire three classes of antiretroviral.

4.5 HIV-1 drug resistance associated Mutations

Four hundred and sixty-four (464) mutations were described. Matayos hospital documented the uppermost over-all (n=196) while the lowest frequency was observed in Khunyangu. Mutations frequency decoding resistance against NNRTIs was the topmost tracked by NRTIs then PIs in that order. Matayos led with mutations encrypting resistance against both NNRTIs and NRTIs. This is because more participants in Matayos were utilization NNRTIs and NRTIs drugs comparable to Busia and Khunyangu. Thus, mutation, M184V which encrypt resistance against 3TC, a prime NRTI was more rampant in Matayos comparable to other two hospitals. More mutations encrypting resistance against PIs were observed in Khunyangu hospital a reason being ATV/r and LPV/r as PIs based drugs were mostly being utilized by many participants in this facility compared to Busia and Matayos hospitals (Table 6).

Table 6: Frequency of class associated Mutations

NRTIs				NNRTIs				PIs			
Mut'	**B'**	**K'**	**M'**	**Mut'**	**B'**	**K'**	**M'**	**Mut'**	**B'**	**K'**	**M'**
D67Y	1	0	0	E138A	2	3	2	I501T	1	0	0
D67N	4	3	4	A98G	1	2	9	L23I	1	0	0
K219E	3	1	2	G190A	5	7	8	150L	1	0	2
E44D	1	0	0	F227L	1	0	0	M461	1	1	3
K65R	9	4	10	K101E	4	5	8	G48V	1	1	0
K219Q	3	0	5	H221Y	3	6	5	F53L	0	2	0
K70N	1	0	0	K103N	19	15	21	I54V	1	4	0
K70E	1	3	4	K101P	3	0	0	K20T	0	1	0
L210W	1	0	0	K238T	1	0	1	V82A	2	2	0
K70R	4	0	4	K103S	2	1	1	L89T	0	1	0
L74L/V	1	0	0	M230L	1	2	1	G48A	0	1	0
L74I	2	1	4	L1001	1	0	4	V82S	0	1	0
M184G	1	0	0	P225H	3	3	7	L10F	0	1	0
L74V	2	2	1	M230MI	1	0	0	L241	0	0	1
M184L	1	0	0	V106A	1	0	0	Q58E	0	1	0
M184I	1	1	2	V1061	1	1	1	V32I	0	0	2
M184V	20	17	31	V179T	2	0	2	I47A	0	0	1
M184M/V	1	0	0	V108I	2	1	3	1461	1	0	0
T215F	4	2	5	Y188C	1	0	0	L33F	0	0	1
M41L	3	1	3	Y181C	6	9	12	V82M	0	0	2
T69G	1	0	1	K101H	0	1	0				
T215Y	1	1	0	G190S	0	1	1				
V75I	2	1	2	V179E	0	1	1				
V106A	1	0	0	V179D	0	1	0				
Y115F	1	5	5	Y181I	0	1	1				
V75M	1	0	0	V179F	0	3	0				
A62V	0	1	0	M230I	0	0	1				
Y181V	1	0	0	Y188L	0	1	3				
T69D	0	1	0	Y181YC	0	0	1				
K70T	0	1	0	V108VI	0	0	1				
D67D	0	0	1	E138G	0	0	1				
V65R	0	1	1								
K219KR	0	0	1								
A62V	0	0	2								
K219R	0	0	1								
	72	46	89		60	64	95		10	16	12

Table legend: Mut'mutations; M', Matayos health facility; K', Khunyangu health facility; B, Busia health facility. Mutations dyed in yellow color were reported in historical studies while in green are reported for the first time in Kenya in this study. Descriptive tests in SPSS v. 20 was utilized to calculate number of mutations. Systematic-reviews was executed by extracting 200 related manuscripts through MEDLINE to describe mutations already reported in the past studies and currently identified for the first time.

CHAPTER FIVE

DISCUSSION

5.1 Socio Demographical and Clinical Characteristics

The present research was planned to describe the HIV type 1 prevalence and sub-types, characterize mutations associated with HIV type 1 drug resistance and patterns amongst participants in the County of Busia, Kenya. Minimum of 200 participants' blood samples with viral counts of ≥ 500copies per microliter of blood were considered for the current study. Plasma samples consisting viral tallies of 1000 copies per microliter of blood gives brilliant results throughout sequencing (Santoro *et al.*, 2014). The present study however, considered plasma samples with viral counts of below 1000 copies per microliter of blood with an aim of minimizing study biasness. Pearson's correlation was utilized to relate two variables, and r value of 0.311 and a $p = 0.028$ between viral tallies and all the mutations, portrayed a robust positive association amid these two variables.

Out of the entire minimum plasma samples of 200, 160, 80% samples (150 ART experienced and 10 inexperienced) were effectively sequenced. This was linked to the context that, the summative samples studied was a representative of the over-all HIV-1 diseased folks in the County of Busia (Ministry of Health and National AIDS Control Council 2018b). However, the residual samples were not sequenced successfully perhaps due to unnoticeable amplified cDNA on the electrophoresed gel. A contradictory outcomes documented report in a study executed in Moshi, Tanzania which had 41.5% samples transiting to sequencing level (Dorothy *et al.*, 2019). Probability of dilapidation of RNA of HIV-1 in the plasma samples during rounds of freezing/thawing might have likewise backed this (Marion *et al.*, 2017). The present study consisted of grownups who were ≥ 18

years and diseased of HIV type 1 and utilizing ART. These groups of persons exist with viruses that harbor mutations encrypting resistance against various antiretroviral.

In the present study, further HIV-1 disease was regularly witnessed amongst females comparable to males endorsing earlier studies (Kinyua *et al.*, 2018, Onywera *et al.*, 2017). This is justified by the fact that HIV type 1 prevalence of females, 5.2% is somewhat more compared to males, 4.5% in Kenya (Ministry of Health and National AIDS Control Council 2018b). Additionally, regardless of males possessing better economic and social leads, they are fewer probable compared to females to pursue health care services, test for HIV infection or commence and observe HIV treatment (United Nations Programme on HIV/AIDS 2017). Supplementary analysis of data showed that, a great portion of grownups of years from 18 to 25 were enrolled in the current study comparable to older counterparts. These outcomes are similar to the outcomes document by the Kenya AIDS Response Progress Report 2016 which described participants of this age set having more HIV-1 prevalence than other age sets (Ministry of Health and National AIDS Control Council 2018a)

Investigation of the participants concerning viral counts revealed that numerous participants possessed viral counts of ≥ 1000 copies per microliter of blood comparable to individuals with viral counts of < 1000 copies per microliter of blood. More frequency of samples with viral counts of less than 1000 copies per microliter might have given un-successfully sequences. Contrary, those which had viral counts of more than 1000 copies per microliter of blood were sequenced efficaciously as it was described in the previous study that, these kind of samples offer paramount outcomes through sequencing (Santoro *et al.,* 2014).

Further analysis of participants in the present study concerning the period of ART uptake specified that more participants had utilized antiretroviral for 5 and more years. This is ascribed by the national regulation that each and every person confirmed positive of HIV-1 be initiated on the therapy irrespective of their CD_4 counts (Ministry of Health and National AIDS & STI Control Program 2018)**.** Lengthier utilization of ART impact the progression of mutations (Aghokeng *et al.,* 2011, Hammer *et al.,* 2011). This conquers with the outcomes in the current study in a manner that, more of the participants who had utilized ARVs for a lengthier period harbored many mutations comparable to those who had utilized ARVs for shorter period of time.

In the current study, partakers were consuming a total of 9 unlike ARVs of drug-regimen combinations that are obtainable for grownups. Supplementary investigation indicated that numerous participants were utilizing first line drug regimen combination, explicitly AZT;$_{3TC}$;NVP, AZT;$_{3TC}$;EFV and TDF;$_{3TC}$;EFV. This is since the previous two drug regimen combinations are superlative first line ART utilized to manage HIV-1 amongst infected individuals and the later regimen is the replacement of the first line ART regimen (Ministry of Health 2018, World Health Organization 2013a, b). These results concordance with the datum that, more participants were utilizing these three drug regimens compared to others. While TDF;$_{3TC}$;NVP is another ideal first line drug regimen combination, it was rather utilized by fewer study participants. This is since TDF which is pooled in TDF;$_{3TC}$;NVP drug regimen combination was initiated for use in the country late compared to AZT and NVP which are pooled in the AZT;$_{3TC}$;NVP and AZT;$_{3TC}$;EFV drug regimens which were introduced earlier and consumed as a sole dose before combination drug regimen for a lengthier period of time. They are also connected with progression of

mutations due to their stumpy genetic wall, sanctioning the advancement of resistance in the gene of reverse transcriptase (Javier *et al.,* 2012) thus proving treatment failure.

TDF;$_{3TC}$;DTG is an ideal first line drug regimen combination in the management of HIV type 1 in Kenya. However, in the current study, it was similarly utilized by fewer participants merely because it is the newest option commended first line drug for utilization in Kenya (National AIDS and STI Control Program 2016). AZT;$_{3TC}$;ATV/r and TDF;$_{3TC}$;ATV/r drug-regimen combinations are ideal 2nd line ARVs and were rather consumed by numerous participants paralleled to AZT;$_{3TC}$;LPV/r and TDF;$_{3TC}$;LPV/r, second line ARVs due to the guiding principle that after dwindling on the above ideal 2nd line drug and both the ideal 1st line and substitute of the 1st line drugs, the participants be changed to AZT_3TC_LPV/r and TDF_3TC_LPV/r which are the substitute of the ideal 2nd line-therapy because numerous patients don't contain LPV/r grounded drug regimen (Ministry of Health 2018, National AIDS and STI Control Program 2016).

5.2 Subtypes of HIV-1

Among the specific goals of this current study was to describe the subtypes of HIV-1 mingling amid community participants of the County of Busia, Kenya. The area of study is situated at the boundary of Kenya and Uganda. It possesses two major points namely Busia and Malaba which aid as passage boundary marker for nomads, crossing to and from Central, East, and West Africa nations with aims of searching better lives, intermarriages and conducting barter trade, sight-seeing and supplementary occupational trade. The study areas entice the introduction of dissimilar subtype of HIV-1 grounding on the listed above explanations.

Sub-type A_1 was the supreme main variant documented in this study. This was considerably followed by sub-type D and sub-type C was infrequently determined. These fallouts are comparable to supplementary studies implemented within the Country which documented the mingling of predominance of sub-type A_1 (Amin *et al.,* 2018, Elizabeth *et al.,* 2015, Elizabeth *et al.,* 2017). Additional past study executed in about 10 years back in similar vicinity of County of Busia documented sub-types D, A_1, and C that were equally described in the current study. Inversely, the present study did not -detect sub-types G, A_2 and B which were documented in the preceding study (Adungo *et al.,* 2014). Other former studies correspondingly documented sub-types D, A_1, and C which were documented in the present study (Oyaro *et al.,* 2011, Otecko *et al.,* 2016, Kageha *et al.,* 2012). One preceding study implemented some years back described a single sub-type of B contrary to the present study (Kamini *et al.,* 2017)

The continual predominance of sub-type A_1 amongst the Kenyan individuals is owed to earlier introduction of this sub-type as documented by its predominance in the preceding studies, varied introduction amongst the research population from various localities of the counties of Kenya and its faster discrepancy and advanced transmissibility relative to other sub-types among the Kenyan populace (Adungo *et al.,* 2014, Elizabeth *et al.,* 2015, Oyaro *et al.,* 2011). Other preceding studies executed in the Country (Adungo *et al.,* 2014, Kamini *et al.,* 2017 Kinyua *et al.,* 2018) documented sub-type D equally the current study. This sub-type predominates in Ugandan populace (Susanna *et al.,* 2020). This is apparent that vicinities at the boundary entice the introduction of unalike sub-types into the populace of Kenya because of the explanations specified above. The research site is sited alongside the

Kenya and Uganda boundary and is pronounced by cross boundary migration causing introduction of sub-type D among the Kenyan people.

Sub-type C which was correspondingly documented in the present study, is the utmost rampant in India, South Africa and Tanzania (Elizabeth *et al.,* 2017, Erik *et al.,* 2017, Matthew *et al.,* 2020). Cross boundary migration of people may have triggered the introduction of this sub-type among the Kenyan people. One study documented that sub-type C is the utmost predominant sub-type mingling in the Northern part of Kenya because of cross boundary transit of people into and from Kenya and Ethiopia through the boundary of North of Kenya (Khamadi *et al.,,* 2005). This shows that cross boundary movement propels the presentation of this sub-type amid the population of Kenyan. The current study similarly documented an aggregate of nine HIV type 1 recombinant forms mingling amongst the research populace. The predominant recombinant was A_1-D this was tracked by A_1-C and A-G in this manner. Mingling recombinants B-C, A_1-D, A_1-A_2, A_1-K, A_1-J, were identified in smaller frequencies. These outcomes are similar to former studies which documented A_1-D as the utmost predominant recombinant (Khoja *et al.,,* 2008, Adungo *et al.,,* 2014, Oyaro *et al.,,* 2011, Otecko *et al.,,* 2016)(Kamini *et al.,* 2017).

Other recombinant forms in this study were identified in small numbers similarly in the previous studies portraying their limited transmission (Kamini *et al.,* 2017, Adungo *et al.,* 2014, Oyaro *et al.,* 2011, Otecko *et al*, 2016, Iweriebor *et al.,* 2012). Viral mingling amongst sub-types A_1, D and C predominance in some nations of Central and East Africa (Hemelaar *et al.,,* 2012) might have ascribed to these mixing. Recombinant A_1-A_2 was similarly identified in the current study. This recombinant was likewise documented in the past study conducted in Western Kenya (Oyaro *et al.,* 2011). This represents its narrow

communication as it is obviously portrayed by the present and the preceding study in that a single isolate was documented. Form of recombinant A_1-C was identified in the present study. This inter-subtype crowned from the spread of already virally mixed recombinant form of A_1-C or throughout reproduction of the HIV type 1 virus. Historical research in Kenya have described A_1-C inter-subtype recombinant in lesser frequencies. This conceivably exposes viral mingling amongst sub-types A and C from East African nations and Southern African and Indian countries respectively. In that pure sub-type C is dominant in Southern Africa Countries and Ethiopia (Mourez, Simon and Plantiera, 2013)(Elizabeth *et al.,* 2017).

The present study documented a single inter-subtype recombinant BC which was equally described in the previous study executed in Coastal zone of Kenya, (Hué *et al.,* 2004). Sub-type B customarily is leading in Japan (Shiino *et al.,* 2020), Australia and America (Bbosa *et al.,* 2019), Europe, and Thailand. Sub-type C is principal in some parts of China (Peipei *et al.,* 2019), Southern-Africa and East Africa (Hemelaar, 2013)(Elizabeth *et al.,* 2017) and rampant amongst HIV diseased homo-sexual males in Europe (Brand *et al.,*, 2012). This inter-subtype recombinant conceivably has been presented amongst the Kenyan populace by viral fraternization and also migration of peoples from these familiar nations into Kenya.

This study documented single inter-subtype recombinant forms of A_1G and A_2D which were equally documented in previous studies conducted in East Africa (Marta *et al.,* 2010) and in Kisumu (Otecko *et al.,* 2016). These recombinants conceivably have been presented into the current research population by viral mixing amongst the residence of county of Busia as well as migration of folks from other locales of Kenya. Another study

implemented in a decade years back in the same locality described sub-types D, A_2, A_1, and G (Adungo *et al.*, 2014), nevertheless, the present study described inter-subtypes recombinants A_2-D and A_1-G which conceivably have occasioned from HIV-1 viral mixing of A_1 and G, A_2 and D amongst the investigated populace.

The present study documents for the initial time inter-subtypes A_1J and A_1K which were formerly not reported in the related research in Kenya (Adungo *et al.*, 2014, Kageha *et al.*, 2012, Kinyua *et al.*, 2018, Oyaro *et al.*, 2011). The inevitability of the spread patterns of the above two inter-subtype recombinants is unidentified since our study was primary to report the existence of these inter-subtypes in Kenyan.

The advancement of inter-subtypes A_1J and A_1K and re-occurrence of A_1-B, A_1-A_2, A_2-D, A_1-D, B-C, A_1-G and A_1-C inter-subtype recombinants entitles that the body of a human being is extra fit for the survival of inter-subtype recombinants (Maria and Carlo 2013) and chiefly through this era of ART coverage scale up. The inter-subtype recombinants could have similarly occasioned by the spread of previously mixed variants from HIV type 1 dually diseased peoples among the researched population or in another place (Marta *et al.*, 2020). It is projected that there will be greater than before range of HIV type 1 subtypes and inter-subtype recombinants amongst the population of Busia.

HIV type 1 sub-types exert diverse effects on the advancement of an ailment, degrees of transmission and HIV type 1 drug resistance progression (Santoro and Perno, 2013, Mourez, Simon and Plantiera, 2013). Previous research documented the transmission of HIV type 1 sub-type C from mothers to their fetuses easily (Renjifo *et al.*,, 2004). Additional study observed that Sub-type C divulges amplified vaginal molting comparable to sub-types D and A (John-Stewart *et al.*, 2005). Further study in the republic of Uganda

reported speedier ailment advancement with sub-type D comparable to subtype A (Kiwanuka *et al.,* 2008). Additional study in the republic of Kenya showed sub-type C encrypting mutations against NNRTIs and NRTIs compared to sub-types D and A (Cissy *et al.,* 2017).

5.3 Prevalence of HIV-1 Drug Resistance

Additional specific aim of the current study was to define the HIV type 1 drug-resistance prevalence. This study observed an increased HIV type 1 drug resistance prevalence among the partakers attending all the three hospitals paralleled to occurrences reported in the former studies (Budambula *et al.,* 2015, Kinyua *et al.,* 2018, Koigi *et al.,* 2014, Lel *et al.,* 2014, Ndembi *et al.,* 2011, Osman *et al.,* 2013, Onywere *et al.,* 2017, Sampathkumar *et al.,* 2014, Hamers *et al.,* 2011, Sigalof *et al.,* 2012) by a greater margin. The World Health Organization policies that entire individual diseased of HIV to be initiated on ARVs treatment irrespective of CD_4 tallies (Ministry of Health 2018) might have contributed to this increased prevalence paralleled to prevalence in the earlier studies before this policies. Supplementary, the scale up ART coverage of 95% in the County of Busia than the National, 75% (Ministry of Health and National AIDS Control Council 2018b) might have equally backed this elevated HIV type 1 drug resistance prevalence in the present study. This findings concordance with another previous study which confirmed the linking of increased prevalence of drug resistant of transmitted HIV-1 with increased ART coverage (Frentz *et al.,* 2012).

The HIV type 1 drug resistance prevalence of entire major mutations was computed at 73.3% among the studied participants on ART. This was parallel to another previous prevalence observed at the National hospital of Kenyatta 23.1% (Kinyua *et al.,* 2018). This

difference was endorsed by the point that; the former study participants were utilizing first line drug regimen combinations while the present were taking both the first/second line drug regimen combinations. Connotation that afore somebody is changed to second-line drug has previously failed on the first-line drugs (Ministry of Health, National AIDS & STI Control Program, 2018). Consequently, the present study observed increased prevalence paralleled to the former study executed amongst drug-users in coastal region of Mombasa 13.8% (Osman *et al.,* 2013). This dissimilarity is ascribed by the point that the earlier study was implemented in about a decade years back of which resistance advance with longer period of exposure to drugs (Aghokeng *et al.,* 2011, Hammer *et al.,* 2011). Additionally, the increased documented prevalence in the current study paralleled the previous (Rency *et al.,* 2014) carried out in the same vicinity, because the past study was executed amongst children of five years and below whose ART coverage is 78% smaller than the grownups, 95% in the present study (Ministry of Health and National AIDS Control Council, 2018). These findings are concordance with the results of another former study which described the association of increased drug resistance prevalence with scale-up of Art-coverage (Frentz *et al.,* 2012).

In contrast of the findings of the present study, the drug-resistance prevalence was compared to prevalence reported by two previous interrelated studies executed amongst drug unexperienced participants in Western Kenya (Onywera *et al.,* 2017) and Rift valley (Kiptoo *et al.,* 2013). Nevertheless, the outcomes of the present study were comparable to other two earlier studies conducted amongst drug experienced partakers whose prevalence of drug resistance was 70% (Hemmers, *et al.,* 2011). This depicts the selection of viral mutations throughout drug exposure (Taylor *et al.,* 2021)

Categorical of HIV type 1 drug resistance prevalence of entire three-class of drugs were identical amongst the participants crosswise three hospitals excluding prevalence of NRTIs in Matayos. The greater incidences of mutations coding resistance against NRTIs in Matayos compared to two facilities backed this raised prevalence in Matayos hospital. Furthermore, a None-Governmental-Organization (NGO), Academic-Model Providing Access to Healthcare (AMPATH) plays a vital role in Khunyangu and Busia hospitals to certify that anti-HIV-1 drugs are obtainable by patients and patients are sustained on antiretroviral by giving food to HIV-1 diseased individuals and their keens men in Busia and Khunyangu (Beryl 2021). This Organization is not established in Matayos hospital.

In the current study, the HIV type 1 drug resistance prevalence of NRTIs and NNRTIs were also higher compared to earlier study executed amongst individuals of five-year and below (Rency *et al.,* 2014). This is linked to longer exposure to antiretroviral by participants as it is obvious by the results in the current study that, more than a half of the participants had utilized ARVs for a period of six or more number of years. This is in tandem with the outcomes in the previous study that lengthier utilization of ARVs may lead to advancement of drug resistance (Hammers *et al.,* 2013, Aghokeng *et al.,* 2011). Similarly, the drug-resistance prevalence of protease inhibitors was greater in the current study compared to earlier studies implemented in coastal region of Kenya and Nairobi (Budambula *et al.,* 2015, Sampathkumar *et al.,*2014). This difference was ascribed by the datum that the former studies involved drug unexperienced participants while the present involved most of the drug-experienced individuals. On the other hand, the current study did not identify any major mutations encoding resistance to any available class of ARVs among the 10 drug-naïve participants, except minor mutations that have little impact on the HIV-1 drug

resistance. This is in concordance with another study which documented that, the mutations advancement is linked with exposure to drugs (Rossouw *et al.,* 2015).

5.4 Pattern Prevalence of HIV-1 Drug Resistance

This study was premeditated to describe the pattern prevalence of HIV-1 drug resistance. Majority of samples harbored mutations encoding resistance to dual class, NNRTIs/NRTIs antiretroviral (National AIDS and STI Control Program 2012). These outcomes are ascribed to the fact that NRTIs and NNRTIs form first line drug regimen combination and have been utilized for longer period of time compared to PIs which form part of the second line drug regimen combination (Günthard *et al.,* 2019). Furthermore, many of the identified mutations encode resistance against Lamivudine which forms the backbone of all the drug regimen combinations (Taylor *et al.,* 2021). This observation is similar to former study which reported more of infected Kenyans were utilizing NVP which forms part of the first line therapy with significant fraction as great as 35% failing treatment and representing poor adherence (Ochieng *et al.,* 2015).

The prevalence of dual class HIV type 1 drug resistance of NRTIs/NNRTIs documented in the present study was more compared to the prevalence determined in the previous study amongst ART inexperienced folks. The dissimilarity is ascribed by viral selection of mutation during utilization of antiretroviral drugs (Rossouw *et al.,* 2015). Comparing the present study with the previous study executed correspondingly amongst ART experienced partakers (Hammer *et al.,* 2011), the dual-class prevalence of NNRTIs/NRTIs was greater due to the fact that the former study was carried out a decade number of years back which is understood that mutation advance with lengthier exposure to ARVs (Aghokeng *et al.,* 2011, Hammer *et al.,* 2013). The patterns prevalence of HIV-1 drug-resistance particularly

for both NNRTIs and NRTIs in the current study was greater paralleled to another alike previous study conducted amongst children of ≤ 5 years (Lel *et al.,* 2014). The difference was attributable by the declined HIV-1 transmission to children from mothers (Pricilla *et al.,* 2018) and lower coverage of ART uptake among the children (78%) compared to the adults (95%) in County of Busia (Moraes *et al.,* 2014).

Documentation of 8% multiclass prevalence of HIV-1 drug resistance amongst studied partakers is of significance apprehension in that, this leads to slow suppression of the viruses or complete no suppression which is verified by our outcomes on viral tallies and mutations. In contrast with the previous studies (Aghokeng *et al.,* 2011, Budambula *et al.,* 2015, Hamers *et al.,* 2011, Kinyua *et al.,* 2018, Koigi *et al.,* 2014, Lel *et al.,* 2014, Lihana *et al.,* 2009, Onywera *et al.,* 2017, Osman *et al.,* 2013, Robert and Shafer 2004, Sigaloff *et al.,* 2012, Steegen *et al.,* 2009), none of the study documented multiclass mutations in the same samples. This is explained by the fact that both the PIs and RTIs drugs are becoming more obtainable in the Country (Fretz *et al.,* 2012) which is backed up by the ART coverage scale up in several counties of Kenya.

The prevalence of HIV-1 drug resistance of first-line drug regimen combination was somewhat greater paralleled to second line. Arithmetical analysis showed no significance difference between HIV-1 drug resistance of the first and second line drug regimen combination. These results in our study indicate that resistance to second line drug regimen combination may increase sooner thus posing a threat to the entire management of HIV-1. The HIV-1 drug resistance prevalence of first line drug regimen combination in the present study also higher compared to the prevalence observed in the former studies in Nairobi and Mombasa (Kinyua *et al.,* 2018, Koigi *et al.,* 2014, Osman *et al.,* 2013). The prevalence of

drug resistance of second line drug regimen combination was also similarly increased in the current study compared to former studies (Budambula *et al.,* 2015, Koigi *et al.,* 2014). The reason for this differences is due to the fact that; the earlier studies were executed in more than a half a decade years back. This reason agrees with the data of published surveys which documented an increasing prevalence of drug resistance as years went by in the Coastal region of Kenya (Hammers *et al.,* 2011; Sigallof *et al.,* 2012), central Kampala Uganda (Ndembi *et al.,* 2011) and central Entebbe Uganda (Ndembi *et al.,* 2008) due to pressure of antiretroviral.

5.5 HIV-1 Drug Resistance Associated Mutations

The study proposed to determine mutations amongst studied partakers. Further investigation of results indicated that, many samples harbored virus with mutations encoding resistance to both NRTIs and NNRTIs, this is attributable by the fact that the two classes of drugs are used as first line drug regimen combination and all infected persons with HIV-1 are commenced on these classes of regimen before they are changed to second line drug regimen combinations which encompass PIs (World Health Organization 2013a, b). A scenario observed in the current study where by many participants had been utilizing first line drug regimen combination.

As documented by other former studies as well as in the present study, NNRTI linked mutations consisted a majority of the described drug-resistant mutations (Melikian *et al.,* 2014, Tambuyzer *et al.,* 2010), with K103N dictating in frequency. The pronounced presence of K103N mutation is attributable to its easy way of transmission. In addition, NNRTIs in general, poses decreased genetic barrier compared to NRTIs and PIs. Furthermore, the probability of NNRTIs persisting for lengthier period, (Sigalof *et al.,*

2012) could also attribute to its dominancy frequency in comparison to its counterparts Moreover, the higher occurrences of NNRTIs is attributable to the earlier utilization of NVP and EFV single therapy for lengthier period of time or short term NVP and or EFV combination in tripartite ART drug regimen combination, which is associated with greater risk of resistance (Sigalof *et al.,* 2012). This is concordance with the current study since more than 70% partakers had been utilizing NVP and EFV drug regimen triple combination compared to others.

Additional NNRTIs association mutations that were frequently identified in the current study but nevertheless in smaller quantities in the previous studies were E138A, H221Y, G190A, Y181C, K101E, P225H (Aghokeng *et al.,* 2011; Budambula *et al.,* 2015; Hamers *et al.,* 2011; Kinyua *et al.,* 2018; Koigi *et al.,* 2014; Rency *et al.,* 2014; Onywera *et al.,* 2017, Osman *et al.,* 2013; Sigaloff *at al.,* 2012). G190A and Y181C raises amongst individuals utilizing ART for lengthier period of time (Yuncong *et al.,* 2014). This is in agreement with our results thus more than 80% of partakers in this study had been on ARVs for 4 - 13 years.

K101E is a nonpolymorphic mutation advanced among patients utilizing NNRTIs (Tambuyzer *et al.,* 2010). It commonly arises in combination with other NNRTI associated mutations (Melikian *et al.,* 2014) a scenario that was seen with the current study. H221Y is also a nonpolymorphic but accessory NNRTI associated mutation that regularly arises in combination with Y181C (Reuman *et al.,* 2010). A case that was observed in the present study thus participants who harbored H221Y likewise had Y181C. P225H is another nonpolymorphic and accessory mutation encoding resistance against EFV that usually exists in combination with K103N (Bacheler *et al.,* 2000). Existence of K103N alongside

P225H result to more than 50 fold decrease in vulnerability against EFV and NVP (Rhee *et al.,* 2004). A scenario that was seen with our results thus all participants with virus harboring P225H had also K103N. E138A is a poly-morphic mutation that causes a resistance prevalence of 1% to 5% depending on the sub-type. It encrypt resistance to RPV and ETR by decreasing their vulnerability (Tambuyzer *et al.,* 2011) with 2-fold.

Additional analysis showed more participants had again more occurrences of NRTIs associated mutations. This is attributable to the fact that the entire study population had been exposed to NRTIs comprising of 3TC. Nevertheless, the most common mutations with NRTIs was M184V which encrypt resistance against 3TC (Bangsberg *et al.,* 2006). In concordance with other studies in East Africa, we similarly detected high occurrences of M184V mutation (Tambuyzer *et al.,* 2010). This is due to the fact that M184V poses low genetic barrier (Tambuyzer *et al.,* 2010). In addition, the higher occurrences of this mutation is backed by the fact that, 3TC forms the backbone in the triple-combined drug-regimens. The amounts of described NRTI associated mutations advocate a change from the predominantly NNRTI to a pooled RTI drug resistance wide-ranging, with a M184V prevalent as a NRTI. M184V is however usually enormously reverting. Studies have reported a low fitness and late spread of M184V mutation.

K65R encrypting resistance against NRTIs was another mutation detected in large numbers in the current study paralleled to previous studies (Aghokeng *et al.,* 2011, Budambula *et al.,* 2015, Hamers *et al.,* 2011, Kinyua *et al.,* 2018, Koigi *et al.,* 2014, Lel *et al.,* 2014, Lihana *et al.,* 2009, Onywera *et al.,* 2017, Osman *et al.,* 2013, Robert and Shafer 2004, Sigaloff *et al.,* 2012, Steegen *et al.,* 2009). K65R mutation was observed in about 20% of participants who had received TDF ones throughout their treatment. Our study proven that

K65R existed in about 13.4% participants in reception of TDF-combined drug. Dissimilarity was ascribed to less sample count that were involved in the present study. Additionally, three distinct point mutations, K65R, Q151M and M184V, might encode increased level of cross resistance across all accessible NRTIs (Tambuyzer *et al.,* 2011).

Virological failure is associated with synergistic assortment of Y181C, K65R, and G190A resistance (Frentz *et al.,* 2012). These previous outcomes conquer with the current finding in that, participants' samples with viruses harboring the three mutations were attending the hospitals with persisting increased viral tallies portraying virological failure. TDF was initiated as a second-generation NRTI into the clinical practice during the year 2001, with supported antiviral action against HIV (Melikian *et al.,* 2014). This second generation NRTI displayed a better antiviral action against infections due to NAMs (Q151M) and TAMs (D67N, K70R and T215Y) (Melikian *et al.,* 2014). Nevertheless, these outcomes of the preceding studies controvert our study in a manner that, the samples of the participants which had the above mentioned mutations had increased persisting viral tallies. The difference could be backed by the fact that, partakers in the present study had extra NNRTIs and NRTIs linked mutations of clinical significance, which might have decreased the action of TDF based drugs. Other NRTI associated mutation identified were E44D, D67N, K70R, M41L, T215Y/F), L210W, V118I, H208Y and V75I that results to cross resistance amongst the NRTIs (Mercelin *et al.,* 2004). The advancement of these mutations among the study partakers was due to lengthier utilization of AZT

Similarly, other NRTIs and NNRTIs linked mutations of clinical importance were identified in lesser frequencies in current and former studies (Aghokeng *et al.,* 2011, Budambula *et al.,* 2015, Hamers *et al.,* 2011, Kinyua *et al.,* 2018, Koigi *et al.,* 2014, Lel *et*

al., 2014, Lihana *et al.,* 2009, Onywera *et al.,* 2017, Osman *et al.,* 2013, Robert and Shafer 2004, Sigaloff *et al.,* 2012, Steegen *et al.,* 2009). Nevertheless, the advancement of newer mutations and re-occurrence of other mutations shows that the body of human is favorable for the existence of transmuted HIV (Lihana *et al.,*, 2009) and principally in this era of scale up of ART. The presence of mutations might have resulted because of ARVs pressure and spread and respreads of already changed virus amongst HIV-1 infected persons (Lihana *et al.,*, 2009). It is hoped that there will be an amplified range of resistant mutations amongst population of Busia County.

PIs associated mutations were identified in small frequencies. Many mutations of M46I encrypting resistance against PIs were observed in the current study paralleled to the previous study (Sampathkumar *et al.,* 2014). This dissimilarity was due to the fact that, the earlier study involved drug unexperienced participants of which mutations might have advanced due to drug pressure amongst the participants in the present study. Paralleling the current study with the preceding study carried out in the County of Busia on the same topic however in children of ≤ five years old, the former study did not observe any protease associated mutations contrasting the present which observed several proteases linked mutations. This is because none of the participants or their mother in the preceding study had been utilizing the second line drug regimen combination (Lel *et al.,* 2014). Additional study executed amongst participants in the coastal region of Kenya identified none major PIs linked mutations except minor PIs linked mutation (Budambula *et al.,* 2015). Further studies executed in Kampala Uganda and Rift valley in Kenya identified a single major PI linked mutation (Kiptoo *et al.,* 2013; Ndembi *et al.,* 2011). Another study implemented in Nairobi identified none PIs linked mutations amongst patients utilizing the second line drug

regimen combination (Koigi *et al.,* 2014). All these outcomes portray that, there is emerging and re-emerging of PIs linked mutations particularly during the current era of scale up of ART.

CHAPTER SIX

CONCLUSION AND RECOMMENDATIONS

6.1 Conclusions

1. The study documents various HIV type 1 subtypes mingling in Busia County, Kenya with A_1K and A1J sub-types being documented for the first time.
2. The study documents increased prevalence of HIV type 1 drug resistance compared to the earlier studies. This consists an overall prevalence of 73.3% with detailed specific prevalence of NNRTIs, NRTIs, and PIs as 68.0%, 58.0%, and 10.7% respectively. The HIV type 1 drug-resistance prevalence of second- and first line drug regimen combination were 18.6% and 70.1% correspondingly.
3. The pattern prevalence of HIV type 1 drug resistance was documented recorded as follows: dual-class NRTIs/NNRTIs, 46.7%; NRTIs/PIs, 0.7%; and NNRTIs/PIs, 0.7%. Multi-class pattern prevalence of HIV type 1 drug resistance NRTIs/NNRTIs /PIs was 8%. Furthermore, the study has documented further mutations that had not been documented before in Kenya. The study observed that several participants had viruses harboring multi-drug, multi class and dual class mutations which impacted either all or two drug regimen combination ineffective thus lessening suppression or entirely no suppression of the virus occasioning virological failure. This info will lead to enhancement in the development of other effective drugs

6.2 Recommendation for active Action

1. Assortment of HIV type 1 sub-types reported in this study pleas for incessant assessment and monitoring of the association between the drug-resistance and sub-

types if victorious management of HIV has to be attained. This plea for advancement of drugs and vaccine design and development.

2. Mutations encrypting resistance against second line drug regimen combinations amongst participants who had never been initiated on these ARVs and the uppermost frequency of mutations comprising dual/multiclass and multiple mutations depict a hazard to the entire management of HIV and serves as primary data in the commencement of the reviewing of the drug regimen combination.
3. The increased prevalence of HIV type 1 drug resistance documented in this study pleas for reinforcement of the health systems, for instance, inaugurating of routine testing services and availing them to diseased HIV-1 individuals prior to commencement and during an alteration of treatment choices, this will minimize mutations advancement. Additionally, follow-up schedules for HIV-1 diseased individuals utilizing ART would add value in ensuring that they are utilizing the drugs suitably, this would similarly minimize mutations advancement. Persistence of increased viral tallies amongst HIV-1 diseased individuals warrants an alteration of therapy choices

6.3 Recommendation for Future Studies

1. Assortment of HIV type 1 sub-types documented in this study please for endless studies on the assessment and monitoring of the correlation between the HIV type 1sub-types and drug-resistances
2. Mutations encrypting resistance against PIs as second-class drug amongst participants who have never been initiated on this class of drug will act as primary data for investigators to come up with an alternative combinations of drug regimen.

3. Increased HIV type 1 drug resistance prevalence in the population of the County of Busia please for duplication of such alike research in other counties with increased HIV-1 prevalence so that to use the results for the aim of budget planning of health matters.

REFERENCES

Abbott Molecular, I.n.c. and Des Plaines, I.U. (2014) Abbott Laboratories. 2014. , Accessed at www.abbottmolecular.com/products/infectious-diseases/realtime-pcr/hiv-1-assay.html.

Adungo, F.O., Gicheru, M.M., Adungo, N.I., Matilu, M.M., Lihana, R.W. and Khamadi, S.A. (2014) Diversity of Human Immunodeficiency Virus Type-1 Subtypes in Western Kenya. World Journal of AIDS 4, 365-372.

Aghokeng, A.F., Kouanfack, C., Laurent, C., Ebong, E., Atem-Tambe, A., Butel, C., Montavon, C., Mpoudi-Ngole, E., Delaporte, E. and Peeters, M. (2011) Scale-up of antiretroviral treatment in sub-Saharan Africa is accompanied by increasing HIV-1 drug resistance mutations in drug-naive patients. AIDS 25(17), 2183-2188.

Alison, C., Shailendra Sawleshwarkar, Rick Varma, Belinda Herring, Kiran Thapa, Dominic Dwyer, Doris Chibo, Nam Nguyen, Karen Hawke, Rodney Ratcliff, Roger Garsia, Anthony Kelleher, David Nolan and The Australian Molecular Epidemiology Network-HIV (2017) A national study of the molecular epidemiology of HIV-1 in Australia 2005–2012. PLoS One 12(5), e0170601.

Amin, S.H., Joakim Esbjörnsson, Elizabeth Wahome, Alexander Thiong'o, George N. Makau, Mathew A. Price and Eduard J. Sanders (2018) HIV-1 subtype diversity, transmission networks and transmitted drug resistance amongst acute and early infected MSM populations from Coastal Kenya. PLoS One 13(12), e0206177.

Azia, I.N., Mukumbang, F.C. and Van, W.B. (2016) Barriers to adherence to antiretroviral treatment in a regional hospital in Vredenburg, Western Cape, South Africa S Afr J HIV Med 17 (1), a476

Bangsberg, D.R., Acosta EP, Gupta R, Guzman D, Riley ED, Harrigan PR, Parkin N and Deeks SG (2006) Adherence-resistance relationships for protease and non-nucleoside reverse transcriptase inhibitors explained by virological fi tness. AIDS 20, 223–231.

Bbosa, N., Kaleebu, P. and Ssemwanga, D. (2019) HIV subtype diversity worldwide. Current Opinion in HIV and AIDS 14(3), 153-160.

Beloukas, A., Psarris, A., Giannelou, P., Kostaki, E., Hatzakis, A. and Paraskevis, D. (2016) Molecular epidemiology of HIV-1 infection in Europe: An overview. Infect Genet Evol 46, 180-189.

Beryl, M. (2021) Partners Launch Busia County UHC Pilot. March 26, 2021. Accessed at https://www.ampathkenya.org/news-blog-feed/2021/3/24/.

Bin, S., Yin Wang, Ruifeng Zhou, Taiyi Jiang, Zhang, H., Zaicun Li, An Liu, Ying Shao, Wei Hua, Tong Zhang, Hao Wu, Shenghua He, Lili Dai and Lijun Sun (2019) Efficacy and Tolerability of Lopinavir/Ritonavir- and Efavirenz-Based Initial Antiretroviral Therapy in HIV-1-Infected Patients in a Tertiary Care Hospital in Beijing, China. Front. Pharmacol. 10(1472), 1-8.

Budambula, V., Musumba, F.O., Webale, M.K., Kahiga, T.M., Ongecha-Owuor, F., Kiarie, J.N., Sowayi, G.A., Ahmed, A.A., Ouma, C. and Were, T. (2015) HIV-1 protease inhibitor drug resistance in Kenyan antiretroviral treatment-naive and -experienced injection drug users and non-drug users. AIDS Research and Therapy 12(27), 1-9.

Burns, R., Borges, J., Blasco, P., Vandenbulcke, A., Mukui, I., Magalasi, D., Lucas, M., Rolanda, M., Birgit, S. and Alison, W. (2019) ‘I saw it as a second chance’: A qualitative exploration of experiences of treatment failure and regimen change

among people living with HIV on second- and third-line antiretroviral therapy in Kenya, Malawi and Mozambique. Glob Public Health [Internet] 0(0), 1–13.

Bussmann, H., de la Hoz Gomez, F., Roels, T.H., Wester, C.W., Bodika, S.M., Moyo, S., Taffa, N., Anderson, M.G., Mine, M., Bile, E.C., Yang, C., Mphoyakgosi, K., Lehotzky, E.A., Mlotshwa, B., Mmelesi, M., Seipone, K., Makhema, M.J., Marlink, R.G., Novitsky, V. and Essex, M. (2011) Prevalence of transmitted HIV drug resistance in Botswana: lessons learned from the HIVDR-Threshold Survey conducted among women presenting for routine antenatal care as part of the 2007 national sentinel survey AIDS Res Hum Retroviruses 27(4), 365-372.

Chad, J.A., Kristin, M.D., Robert, L.M. and Christine, K. (2011) Atazanavir/ritonavir-based combination antiretroviral therapy for treatment of HIV-1 infection in adults. Future Virol. 6(2), 157-177.

Cissy, K., Jennifer, T., Immaculate, N., Anne, H., , , Emmanuel, N., Colin, W., Ivan, M., Joep, J.O., Wools-Kaloustian, K., Silvia, B., Philippa, J.E., Peter, M., , Sarah, A.W. and Nicholas, I.P. (2017) HIV Drug Resistance Mutations in Non-B Subtypes After Prolonged Virological Failure on NNRTI-Based First-Line Regimens in Sub-Saharan Africa. J Acquir Immune Defic Syndr 1(175), e45-e54.

Cuevas, J.M., Geller R, Garijo R, López-Aldeguer J and Sanjuán R (2015) Extremely High Mutation Rate of HIV-1 In Vivo. PLOS Biology 13(9), e1002251.

Deogratius, S., Nicholas, B., Rebecca, N.N., Alfred, S., Anne, K., Maria, N., Faridah, N., Alex, K., Joseph, M., Janet, S., Gonzalo, Y., Andrew, L.B. and Pontiano, K. (2020) The Molecular Epidemiology and Transmission Dynamics of HIV Type 1 in a General Population Cohort in Uganda. Virus 12 (1283).

Désiré, N., Lorenzo Cerutti, Quentin Le Hingrat, Marine Perrier, Stefan Emler, Vincent Calvez, Diane Descamps, Anne Geneviève Marcelin, Stéphane Hué and Benoit Visseaux (2018) Characterization update of HIV-1 M subtypes diversity and proposal for subtypes A and D sub-subtypes reclassification. Retrovirology 15(80), 1-7.

DHHS (2021) Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the Use of Antiretroviral Agents in Adults and Adolescents

with HIV. Department of Health and Human Services. Available at https://clinicalinfo.hiv.gov/sites/default/files/guidelines/documents/.

Diallo, K., Gotte, M. and Wainberg, M.A. (2003) Molecular impact of the M184V mutation in human immunodeficiency virus type 1 reverse transcriptase Antimicrobial Agents and Chemotherapy. 47(11), 3377-3383.

Dorothy, E.D., Aisa M. Shayo, Coleen K. Cunningham and Blandina Mmbaga (2019) HIV-1 Drug Resistance and Virologic Outcomes Among Tanzanian Youth living with HIV Pediatr Infect Dis 38(6), 617–619.

Elizabeth, L., Rebecca Waihenya, James Munyao, Lucy Sanguli, Christina Mwachari and Samoel Khamadi (2015) HIV-1 Drug Resistance Mutations in Patients Failing 1st Line Therapy in a Comprehensive Care Center in Nairobi, Kenya. World Journal of AIDS 5, 83-89.

Elizabeth, M.E., Lufuno Mavhandu, Cecile Manhaeve, Keanan McGonigle, Patrick Jackson, David Rekosh, Marie-Louise Hammarskjold, Pascal O. Bessong and Denis M. Tebit (2017) High level of HIV-1 drug resistance mutations in patients with unsuppressed viral loads in rural northern South Africa. AIDS Res Ther. 14(36).

Eric, J.A. and Daria, J.H. (2012) HIV-1 antiretroviral drug therapy. Cold Spring Harb Perspect Med 2(4), a007161.

Erik, B., Eric Sanders-Buell, Meera Bose, Gustavo H. Kijak, Andrea Bradfield, Jacqueline Crossler, Miguel A. Arroyo, Leonard Maboko, Oliver Hoffmann, Steffen Geis, Deborah L. Birx, Jerome H. Kim, Nelson L. Michael, Merlin L. Robb, Michael Hoelscher and Sodsai Tovanabutra (2017) HIV-1 Genetic Diversity Among Incident Infections in Mbeya, Tanzania. AIDS Research and Human Retroviruses 33(4), 373-381.

Eungi, C., Jayapal, R.M., Dai, L. and Srikanth, K. (2018) Recent advances in the discovery of small-molecule inhibitors of HIV-1 integrase. Future Sci OA 4(9), FSO338.

Frentz, D., Boucher, C.A. and van de Vijver, D.A. (2012) Temporal changes in the epidemiology of transmission of drug-resistant HIV-1 across the world. AIDS Rev 14(1), 17-27.

Günthard, H.F., Calvez, V., Paredes, R., Pillay, D., Shafer, R.W., Wensing, A.M., Jacobsen, D.M. and Richman, D.D. (2019) Human Immunodeficiency Virus Drug Resistance: 2018 Recommendations of the International Antiviral Society-USA Panel. Clin Infect Dis 68(2), 177-187.

Hamers, R.L., Wallis, C.L., Kityo, C., Siwale, M., Mandaliya, K., Conradie, F., Botes, M.E., Wellington, M., Osibogun, A., Sigaloff, K.C., Nankya, I., Schuurman, R., Wit, F.W., Stevens, W.S., van Vugt, M., de Wit, T.F. and PharmAccess African Studies to Evaluate, R. (2011) HIV-1 drug resistance in antiretroviral-naive individuals in sub-Saharan Africa after rollout of antiretroviral therapy: a multicentre observational study. Lancet Infectious Diseases 11(10), 750-759.

Hamlyn, E., Fiona, M.E., Kholoud P., David A. C., Giuseppe T., Mauro, S., Court, P., Jason F. O., Myra M., Abdel B., Jonathan W. and Sarah, F. (2012) Plasma HIV Viral Rebound following Protocol-Indicated Cessation of ART Commenced in Primary and Chronic HIV Infection. PLoS One 7(8), e43754.

Hassan, A.S., Mwaringa, S.M., Obonyo, C.A., Nabwera, H.M., Sanders, E.J., Rinke de Wit, T.F., Cane, P.A. and Berkley, J.A. (2013) Low prevalence of transmitted HIV type 1 drug resistance among antiretroviral-naive adults in a rural HIV clinic in Kenya. AIDS Res Hum Retroviruses 29(1), 129-135.

Hassan, A.S., Nabwera, H.M., Mwaringa, S.M., Obonyo, C.A., Sanders, E.J., Rinke de Wit, T.F., Cane, P.A. and Berkley, J.A. (2014) HIV-1 virologic failure and acquired drug resistance among first-line antiretroviral experienced adults at a rural HIV clinic in coastal Kenya: a cross-sectional study. AIDS Res Ther 11(1), 9.

Heestermans, T., Browne, J.L., Aitken, S.C., Vervoort, S.C. and Klipstein-Grobusch, K. (2016) Determinants of adherence to antiretroviral therapy among HIV-positive adults in sub-Saharan Africa: a systematic review. BMJ Glob Health Dec 30; 1, e000125.

Ireen, E.K., Vladimir, N., Lauren, M., Jeannie, B., Rachel, M., Noel, S., John, S., Mary, F.M., Saidi, H.K. and Essex, M. (2012) HIV-1 Subtypes and Recombinants in Northern Tanzania: Distribution of Viral Quasispecies. PLoS One 7(10), e47605.

Isabelle, M., Frédéric Subra, Charlotte Charpentier, Gilles Collin, Diane Descamps, Vincent Calvez, Anne-Geneviève Marcelin and Olivier Delelisc (2017) Mutations Located outside the Integrase Gene Can Confer Resistance to HIV-1 Integrase Strand Transfer Inhibitors. mBio 8(5), e00922-00917.

Ivan, N., Ivan, K., Billy, N.M., Patrick, K., Joseph, L., Fred, L., Anne, A.K., Pontiano, K. and Paula, M. (2016) From antiretroviral therapy access to provision of third line regimens: evidence of HIV Drug resistance mutations to first and second line regimens among Ugandan adults BMC Res Notes 9(915).

Javier, E., Concepción, A., Conxa, B. and Francisco, P. (2012) Pharmacological and clinical evidence of nevirapine immediate- and extended-release formulations. HIV AIDS (Auckl) 4, 169–179.

Jeannette, M.W., Neil, T.P., Colombe, C., Nicholas, S.H. and Christos, J.P. (2003) Broad nucleoside reverse-transcriptase inhibitor cross-resistance in human immunodeficiency virus type 1 clinical isolates. The Journal of Infectious Diseases 188(7), 992-1000.

Jones, J., Sullivan, P.S. and Curran, J.W. (2019) Progress in the HIV epidemic: Identifying goals and measuring success. PLOS Medicine 16(1), e1002729.

Jordon, M.R., Bennet DE, Wainberg MA, Havlir D, Hammer S, Yang C, Morris L, Peeters M, Wensing AM, Parkin N, Nachega JB, Phillips A, De Luca A, Geng E, Calmy A, Raizes E, Sandstrom P, Archibald CP, Perriëns J, McClure CM, Hong SY, McMahon JH, Dedes N, Sutherland D and Bertagnolio S (2012) Update on WHO HIV drug resistance prevention and assessment strategy 204-2011. Clin Infect Dis 54(supp 4), 1-13.

Joris, H. (2012) The origin and diversity of the HIV-1 pandemic Trends Mol Med 18(3), 182-192.

Joris, H., Ramyiadarsini, E., Jason, Y., Leslie, D., Isabella, F., Shona, K., Brian, W., Gouws-Williams, E. and Peter, D.G. (2019) Global and regional molecular

epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis. Lancet Infect Dis 18, 143-155.

José, R.A., Emilio, G.Q., Belén, S. and Almudena, G. (2008) Tenofovir: pharmacology and interactions. Enferm Infecc Microbiol Clin 26(Suppl 8), 2-6.

Julie, S.E. and Shivaraj, N. (2021) Highly Active Antiretroviral Therapy (HAART) StatPearls Publishing LLC. last Update: July 8, 2021.

Kageha, S., Lihana, R.W., Okoth, V., Mwau, M., Okoth, F.A., Songok, E.M., Ngaira, J.M. and Khamadi, S.A. (2012) HIV type 1 subtype surveillance in central Kenya. AIDS Res Hum Retroviruses 28(2), 228-231.

Kamini, G., Micah Oyaro, Nagavelli Padayachi, Thando Mbali Zulu, Tulio de Oliveira, John Wylie and Thumbi Ndung'u (2017) Complex Subtype Diversity of HIV-1 Among Drug Users in Major Kenyan Cities. AIDS Res Hum Retroviruses 33(5), 500-510.

Kantor, R., DeLong, A., Balamane, M., Schreier, L., Lloyd, R.M., Jr., Injera, W., Kamle, L., Mambo, F., Muyonga, S., Katzenstein, D., Hogan, J., Buziba, N. and Diero, L. (2014) HIV diversity and drug resistance from plasma and non-plasma analytes in a large treatment programme in western Kenya. J Int AIDS Soc 17, 19262.

Kenya National Bureau of Statistics (2019) Kenya Population and Housing Census: Volume I, population by County and Sub-County. Accessed at: http://housingfinanceafrica.org/app/uploads/VOLUME-I-KPHC-2019.pdf.

Khamadi, S.A., Washington Ochieng, Raphael W Lihana, Joyceline Kinyua, Joseph Muriuki, Joseph Mwangi, Raphael Lwembe, Michael Kiptoo, Saida Osman, Nancy Lagat, Roger Pelle, Anne Muigai, Jane Y Carter, Isao Oishi, Hiroshi Ichimura, D L

Mwaniki, Fredrick A Okoth, Solomon Mpoke and Elijah M Songok (2005) HIV type 1 subtypes in circulation in northern Kenya. AIDS Res Hum Retroviruses 21(9), 810-814.

Khoja, S., Peter Ojwang, Saeed Khan, Nancy Okinda, Reena Harania and Syed ali (2008) Genetic analysis of hiv-1 subtypes in Nairobi, Kenya. Plos one 3(9), e3191.

Kinyua, J.G., Lihana, R.W., Kiptoo, M., Muasya, T., Odera, I., Muiruri, P. and Songok, E.M. (2018) Antiretroviral resistance among HIV-1 patients on first-line therapy attending a comprehensive care clinic in Kenyatta National Hospital, Kenya: a retrospective analysis. PanAfrica Medical Journal 29(1), 1-6.

Kiptoo, M., Brooks, J., Lihana, R.W., Sandstrom, P., Ng'ang'a, Z., Kinyua, J., Lagat, N., Okoth, F. and Songok, E.M. (2013) HIV-1 drug resistance-associated mutations among HIV-1 infected drug-naive antenatal clinic attendees in rural Kenya. BMC Infect Dis 13, 517.

Koigi, P., Ngayo, M.O., Khamadi, S., Ngugi, C. and Nyamache, A.K. (2014) HIV type 1 drug resistance patterns among patients failing first and second line antiretroviral therapy in Nairobi, Kenya. BMC Res Notes 7(1), 890.

Kothari, C.R. (2004) Research Methodology: Methods and Techniques. 2nd Edition. New Age International Publishers, New Delhi.

Kumar, S., Stecher, G. and Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Molecular Biology and Evolution 33, 1870-1874.

Lel, R., Ngaira, J., Lihana, R. and Khamadi, S. (2014) HIV-1 drug resistance mutations among infants born to HIV-positive mothers in Busia, Kenya. AIDS Research Human Retroviruses 30(12), 1236-1238.

Liangqun, H. and Chaoping, C. (2013) Understanding HIV-1 protease autoprocessing for novel therapeutic development. Future Med Chem. 5(11), 1-22.

Lihana, R.W., Khamadi, S.A., Lubano, K., Lwembe, R., Kiptoo, M.K., Lagat, N., Kinyua, J.G., Okoth, F.A., Songok, E.M., Makokha, E.P. and Ichimura, H. (2009) HIV type 1 subtype diversity and drug resistance among HIV type 1-infected Kenyan patients initiating antiretroviral therapy. AIDS Res Hum Retroviruses 25(12), 1211-1217.

Lisa, R., Michael, L.L., Qiming, L., Brian, W., Allan, E.R., Winkler, W. and Mark, S. (2015) Prevalence of Antiretroviral Drug Resistance and Resistance-Associated Mutations in Antiretroviral Therapy-Naïve HIV-Infected Individuals from 40 United States Cities. HIV Clinical Trials 8(1), 1-8.

Margolis, D.A., Brinson, C.C., Smith, G.H., de Vente, J., Hagins, D.P., Eron, J.J., Sandy, K.G., Marty, H.S., Marita C. S., Peter, E.W., Susan, L.F., Britt, S.S., Melinda, M.B., Krischan, J.H., Kimberly, Y.S. and William, R.S. (2015) Cabotegravir plus rilpivirine, once a day, after induction with cabotegravir plus nucleoside reverse transcriptase inhibitors in antiretroviral-naive adults with HIV-1 infection (LATTE): a randomised, phase 2b, dose-ranging trial. Lancet Infect Dis. Lancet Infect Dis 15(10), 1145-1155.

Maria, G.Q., Benedetta Mattioli and Stefano Vella (2012) Glances in Immunology of HIV and HCV Infection. Advances in Virology Volume 2012(434036), 1-14.

Maria, M.S. and Carlo, F.P. (2013) HIV-1 Genetic Variability and Clinical Implications. ISRN Microbiol 481314.

Marion, C., Astrid Gall, Monique Vink, FoklaZorgdrager, Špela Binter, Stephanie Edwards, Suzanne Jurriaans, Margreet Bakke, Swee HoeOng, Luuk Gras, Ardvan Sighem, Daniela Bezemer, Peter Reiss, Paul Kellam, Ben Berkhout, Christophe Fraser, Antoinette C.van der Kuyl and BEEHIVE Consortium (2017) From clinical sample to complete genome: Comparing methods for the extraction of HIV-1 RNA for high-throughput deep sequencing. Virus Research 239(15 July), 10-16.

Marta, G., Massimo, C., Cristina, P. and Alessandra, B. (2020) Molecular Epidemiology of HIV-1 in African Countries: A Comprehensive Overview. Pathogens 9(12), 1072.

Matthew, J.G., Michael Roche, Melissa J. Churchill, Paul R. Gorry and Jacqueline K. Flynn (2020) Understanding the mechanisms driving the spread of subtype C HIV-1. EbioMedicine 53, 102682.

McColl, D.J. and Chen, X. (2010) Strand transfer inhibitors of HIV-1 integrase: Bringing IN a new era of antiretroviral therapy. Antiviral Res 85: 101–118Miller V 2001. International perspectives on antiretroviral resistance. Resistance to protease inhibitors. J Acquir Immune Defic Syndr 26 (Suppl 1), S34–S50.

Melikian, G.L., Rhee, S.Y., Varghese, V., Porter, D., White, K., Taylor, J., Towner W, Troia P, Burack J, Dejesus E, Robbins GK, Razzeca K, Kagan R, Liu TF, Fessel WJ, D, I. and RW, S. (2014) Non-nucleoside reverse transcriptase inhibitor (NNRTI) cross-resistance: implications for preclinical evaluation of novel NNRTIs and clinical genotypic resistance testing. J Antimicrob Chemother 2014 69(1).

Menéndez-Arias, L. (2011) A structural frame for understanding the role of thymidine analogue resistance mutations in resistance to zidovudine and other nucleoside analogues. Antivir Ther 16(7), 943-946.

Meriki, H.D., Tufon, K.A., Anong, D.N., Atanga, P.N., Anyangwe, I.A., Fidelis Cho-Ngwa and Theresa Nkuo-Akenji (2019) Genetic diversity and antiretroviral resistance-associated mutation profile of treated and naive HIV-1 infected patients from the Northwest and Southwest regions of Cameroon. . PLoS One 14(11), e0225575.

Michele, W.T. and Robert, W.S. (2012) HIV-1 Antiretroviral Resistance, Scientific Principles and Clinical Applications. Drugs 72(9), e1–e25.

Min, H.K., Je Eun Song, Jin Young Ahn, Yong Chan Kim, Dong Hyun Oh, Heun Choi, Hea Won Ann, Jae Kyoung Kim, Sun Bean Kim, Su Jin Jeong Nam Su Ku, Sang Hoon Han, Young Goo Song and June Myung Kim (2013) HIV Antiretroviral Resistance Mutations Among Antiretroviral Treatment-Naive and -Experienced Patients in South Korea. AIDS Res Hum Retroviruses 29(12), 1617-1620.

Ming-Tain, L., Vandna Munshi, Meiqing Lu, MeiZhen Feng, Renee Hrin-Solt, Philip M. McKenna, Daria J. Hazuda and Michael D. Miller (2016) Mechanistic Study of Common Non-Nucleoside Reverse Transcriptase Inhibitor-Resistant Mutations with K103N and Y181C Substitutions. Viruses 8(10), 263.

Ministry of Health (2018) Guidelines on the use of antiretroviral drugs for treating and preventing HIV infection in Kenya

Ministry of Health (2020) Kenya's National HIV Survey Shows Progress Towards Control of the Epidemic.

Ministry of Health and National AIDS & STI Control Program (2018) Guidelines on the use of antiretroviral drugs for treating and preventing HIV infection in Kenya. 2018 Edition.

Ministry of Health and National AIDS Control Council (2018a) Kenya AIDS Response Progress Report 2018. https://nacc.or.ke/wp-content/uploads/2018/2011/KARPR-Report_2018.pdf.

Ministry of Health and National AIDS Control Council (2018b) Kenya HIV estimates.

Minjie, C., Wuhong Zhang, Xuan Zhang, Wenjie Jiang, Xiping Huan, Xiaojun Meng, Bowen Zhu1, Yue Yang, Yusha Tao, Tian Tian, Yihua Lu, Liying Jiang, Lei Zhang and Xun Zhuang (2017) HIV-1 CRF01_AE strain is associated with faster HIV/AIDS progression in Jiangsu Province, China. Sci Rep 7(1570).

Moraes, D.C., Oliveira, R.C. and Costa, S.F. (2014) Adherence of men living with HIV/AIDS to antiretroviral treatment. Anna Nery School Journal of Nursing 18(4), 676-681.

National AIDS and STI Control Program (2011) Guidelines for antiretroviral therapy in Kenya 4th edition 2011

National AIDS and STI Control Program (2012) Guidelines for antiretroviral therapy in Kenya, 4th edn. 2011.

National AIDS and STI Control Program, E. (2016) Guidelines on Use of Antiretroviral Drugs for Treating and Preventing HIV Infections in Kenya. 2016 edition. Accessed at: www.nascop.or.ke/index.php/new-guidelines/.

National AIDs Control Council (2018a) Homa Bay County Report on The HIV Implementing partners online reporting system (HIPORS) for the financial year 2016/2017, FEBRUARY 2018.

National AIDs Control Council (2018b) Migori County Report on The HIV Implementing partners online reporting system (HIPORS) for the financial year 2016/2017, FEBRUARY 2018.

National AIDs Control Council (2018c) Siaya County Report on The HIV Implementing partners online reporting system (HIPORS) for the financial year 2016/2017, FEBRUARY 2018.

Ndembi, N., Hamers, R.L., Sigaloff, K.C., Lyagoba, F., Magambo, B., Nanteza, B., Watera, C., Kaleebu, P. and Rinke de Wit, T.F. (2011) Transmitted antiretroviral drug resistance among newly HIV-1 diagnosed young individuals in Kampala. AIDS 25(7), 905-910.

Ndembi, N., Lyagoba, F., Nanteza, B., Kushemererwa, G., Serwanga, J., Katongole-Mbidde, E., Grosskurth, H. and Kaleebu, P. (2008) Transmitted Antiretroviral Drug Resistance Surveillance among Newly HIV Type 1-Diagnosed Women Attending an Antenatal Clinic in Entebbe, Uganda. AIDS Res Hum Retroviruses 24(6), 889–895.

Nicolas, A.M., Pamela, W., Rima, K., Kirsten, W., Danielle, P., Michael, E.A., Christian, C. and Michael, D.M. (2017) Commonly Transmitted HIV-1 Drug Resistance Mutations in Reverse-Transcriptase and Protease in Antiretroviral Treatment–Naive Patients and Response to Regimens Containing Tenofovir Disoproxil Fumarate or Tenofovir Alafenamide. The Journal of Infectious Diseases 215(6), 920–927.

Nikolenko, G.N., Delviks-Frankenberry, K.A. and Pathak, V.K. (2010) A novel molecular mechanism of dual resistance to nucleoside and nonnucleoside reverse transcriptase inhibitors. J Virol May;84(10), 5238-5249.

Ochieng, W., Kitawi, R.C., Nzomo, T.J., Mwatelah, R.S., Kimulwo, M.J., Ochieng, D.J., Kinyua, J., Lagat, N., Onyango, K.O., Lwembe, R.M., Mwamburi, M., Ogutu, B.R., Oloo, F.A. and Aman, R. (2015) Implementation and Operational Research: Correlates of Adherence and Treatment Failure Among Kenyan Patients on Long-term Highly Active Antiretroviral Therapy. Journal of Acquired Immune Deficiency Syndrome 69(2), e49-56.

Olipher, M., David Hughes Mulama, John Muoma and Matilu Mwau (2021) Correlation of HIV-1 drug resistant mutations and virologic failure. Pan Afr Med J. 39(180).

Olipher, M., Maureen, A., David, H.M., John, M., Ferdinard, A., Humphrey, W., Asiko, O., Geoffrey, M.M. and Matilu, M. (2020) Prevalence of human immunodeficiency virus-1 drug-resistant mutations among adults on first- and second-line antiretroviral therapy in a resource-limited health facility in Busia County, Kenya. Pan African Medical Journal 37(311).

Onyango, D.O., van der Sande., Musingila P., Kinywa E., Opollo V., Oyaro B., Nyakeriga E., Waruru A, Wanjiru W., Mwangome M., Macharia T., Peter W., Muthoni J., Ngugi C., Kevin M.D. and George W.R. (2019) High HIV prevalence among decedents received by two high-volume mortuaries in Kisumu, Western Kenya PLoS ONE 16(Sterne JA), e0253516.

Onywera, H., Maman, D., Inzaule, S., Auma, E., Were, K., Fredrick, H., Owiti, P., Opollo, V., Etard, J.F., Mukui, I., Kim, A.A. and Zeh, C. (2017) Surveillance of HIV-1 pol

transmitted drug resistance in acutely and recently infected antiretroviral drug-naive persons in rural western Kenya. PLoS One 12(2), e0171124.

Osman, S., Lihana, R.W., Kibaya, R.M., Ishizaki, A., Bi, X., Okoth, F.A., Ichimura, H. and Lwembe, R.M. (2013) Diversity of HIV type 1 and drug resistance mutations among injecting drug users in Kenya. AIDS Res Hum Retroviruses 29(1), 187-190.

Otecko, N., Seth Inzaule, Collins Odhiambo, George Otieno, Valarie Opollo, Alex Morwabe, Kennedy Were, Kenneth Ndiege, Fredrick Otieno, Andrea A. K and Clement Zeh (2016) Viral and host characteristics of recent and established HIV-1 infections in Kisumu based on a Multiassay approach. Scientific Reports 6(37964), 1-9.

Oyaro, M., Mbithi, J., Oyugi, F., Laten, A., Anzala, O. and Engelbrecht, S. (2011) Molecular Characterization of HIV type 1 among HIV infected respondents in a cohort being prepared for HIV phase III vaccine clinical trials, Western Kenya Aids Research and Human Retroviruses 27(3), 357-264.

Parboosing, R., Naidoo, A., Gordon, M., Taylor, M. and Vella, V. (2011) Resistance to antiretroviral drugs in newly diagnosed, young treatment-naive HIV-positive pregnant women in the province of KwaZulu-Natal, South Africa. J Med Virol 83(9), 1508-1513.

Peipei, X., Ying Zhou, Jing Lu, Li Yan, Xiaoqin Xu, Haiyang Hu, Jianjun Li, Ping Ding, Tao Qiu, Gengfeng Fu, Xiping Huan and Haitao Yang (2019) HIV-1 genotype diversity and distribution characteristics among heterosexually transmitted population in Jiangsu province, China. Virology Journal 16(51).

Pleuni, S.P. (2013) HIV Drug Resistance: Problems and Perspectives. Infect Dis Rep 5(Suppl 1), e5.

Price, M.A., Wallis, C.L., Lakhi, S., Karita, E., Kamali, A., Anzala, O., Sanders, E.J., Bekker, L.G., Twesigye, R., Hunter, E., Kaleebu, P., Kayitenkore, K., Allen, S., Ruzagira, E., Mwangome, M., Mutua, G., Amornkul, P.N., Stevens, G., Pond, S.L., Schaefer, M., Papathanasopoulos, M.A., Stevens, W., Gilmour, J. and Group, I.E.I.C.S. (2011) Transmitted HIV type 1 drug resistance among individuals with recent HIV infection in East and Southern Africa. AIDS Res Hum Retroviruses 27(1), 5-12.

Pricilla, R.A., Brown, M., Wexler, C., Maloba, M., Gautney, B.J. and Finocchario-Kessler, S. (2018) Progress Toward Eliminating Mother to Child Transmission of HIV in Kenya: Review of Treatment Guidelines Uptake and Pediatric Transmission Between 2013 and 2016-A Follow Up. Matern Child Health J 22(12), 1685-1692.

Quashie, P.K., Mesplede, T. and Wainberg, M.A. (2013) HIV drug resistance and the advent of integrase inhibitors. Curr Infect Dis Rep 15(1), 85–100.

Radmila, H., Masoud, T. and Bin, T. (2017) RNA-Seq methods for transcriptome analysis. Wiley Interdiscip Rev RNA 8(1), 10.1002/wrna.1364.

Reuman, E.C., Rhee SY, Holmes SP and Shafer RW (2010) Constrained patterns of covariation and clustering of HIV-1 non-nucleoside reverse transcriptase inhibitor resistance mutations. J Antimicrob Chemother 65(7), 1477-1485.

Rhee, S.Y., Liu, T., Ravela, J., Gonzales, M.J. and Shafer, R.W. (2004) Distribution of human immunodeficiency virus type 1 protease and reverse transcriptase mutation

patterns in 4,183 persons undergoing genotypic resistance testing. Antimicrob Agents Chemother 48(8), 3122-3126.

Rhee, S.Y., Varghese V, Holmes SP, Van Zyl GU, Steegen K, Boyd MA, Cooper DA, Nsanzimana S, Saravanan S, Charpentier C, de Oliveira T, Etiebet MA, Garcia F, Goedhals D, Gomes P, Günthard HF, Hamers RL, Hoffmann CJ, Hunt G, Jiamsakul A, Kaleebu P, Kanki P, Kantor R, Kerschberger B, Marconi VC, D'amour Ndahimana J, Ndembi N, Ngo-Giang-Huong N, Rokx C, Santoro MM, Schapiro JM, Schmidt D, Seu L, Sigaloff KCE, Sirivichayakul S, Skhosana L, Sunpath H, Tang M, Yang C, Carmona S, Gupta RK and Shafer RW (2017) Mutational Correlates of Virological Failure in Individuals Receiving a WHO-Recommended Tenofovir-Containing First-Line Regimen: An International Collaboration. EBioMedicine 18, 225-235.

Robert, W. and Shafer, M. (2004) Genotypic Testing for HIV-1 Drug Resistance. Knowl Base Chapter HIV Insite.

Robert, W.S., Kathryn, D., Mark, A.W. and Susan, H.E. (2001) A Guide to HIV-1 Reverse Transcriptase and Protease Sequencing for Drug Resistance Studies. HIV Seq Compend 2001, 1-51.

Robertson, S.M., Penzak, S.R., Lane, J., Pau, A.K. and Mican, J.M. (2005) A potentially significant interaction between efavirenz and phenytoin: a case report and review of the literature Clin Infect Dis Jul 15, 41(2), e15-18.

Robyn, T., Mark, G. and Christopher, J.D. (2011) Lopinavir/Ritonavir: A Review Clinical Medicine Insights: Therapeutics 2011:3 93–102 3 93–102.

Rogers, K. (2008) AZT. Encyclopedia Britannica https://www.britannica. com/science /AZT. Accessed 9 September 2021.

Rossouw, T.M., Feucht, U.D., Melikian, G., van Dyk, G., Thomas, W., du Plessis, N.M. and Theunis, A. (2015) Factors Associated with the Development of Drug Resistance Mutations in HIV-1 Infected Children Failing Protease Inhibitor-Based Antiretroviral Therapy in South Africa. PLoS ONE 10(7), e0133452.

Sampathkumar, R., Shadabi E, La D, John Ho, Binhua Liang, Joshua Kimani, Terry B Ball, Francis A Plummer and Ma Luo (2014) Naturally occurring protease inhibitor resistance mutations and their frequencies in HIV proviral sequences of drug-naïve sex workers in Nairobi, Kenya. Retrovirology 11(Suppl 1), P133. doi:110.1186/1742-4690-1111-s1181-p1133

Santoro, M.M., Lavinia Fabeni, Daniele Armenia, Claudia Alteri, Domenico Di Pinto, Federica Forbici, Ada Bertoli, Domenico Di Carlo, Caterina Gori, Stefania Carta, Valentina Fedele, Roberta D'Arrigo, Giulia Berno, Adriana Ammassari, Carmela Pinnetti, Emanuele Nicastri, Alessandra Latini, Chiara Tommasi, Evangelo Boumis, Nicola Petrosillo, Gianpiero D'Offizi, Massimo Andreoni, Francesca Ceccherini-Silberstein, Andrea Antinori and Carlo Federico Perno (2014) Reliability and clinical relevance of the HIV-1 drug resistance test in patients with low viremia levels Clinical Infectious Diseases 58(8), 1156–1164.

Santosh, K., Sourav, S. and Vangal K. S. (2019) Absence of Integrase Strand Transfer Inhibitor Associated Resistance in Antiretroviral Therapy Naïve and Experienced Individuals from Western India. AIDS Res Hum Retroviruses 35(6), 567-571.

Schultz, A.K., Ming Zhang, Ingo Bulla, Thomas Leitner, Bette Korber, Burkhard Morgenstern and Mario Stanke (2009) JPHMM: Improving the reliability of recombination prediction in HIV-1. Nucleic Acids Research 37(sup.2), 647-651.

Shabani, R.M., Happiness, H.K., Hellen, B.A., Tolbert, B.S., Michael, J.M., Sia, E.M. and Ireen, E.K. (2020) Patterns of acquired HIV-1 drug resistance mutations and predictors of virological failure in Moshi, Northern Tanzania. PLoS One 15(9), e0232649.

Shiino, T., Hachiya, A., Hattori, J., Sugiura, W. and Yoshimura, K. (2020) Nation-Wide Viral Sequence Analysis of HIV-1 Subtype B Epidemic in 2003–2012 Revealed a Contribution of Men Who Have Sex With Men to the Transmission Cluster Formation and Growth in Japan. Front. Reprod. Health 2(531212).

Sigaloff, K.C., Mandaliya, K., Hamers, R.L., Otieno, F., Jao, I.M., Lyagoba, F., Magambo, B., Kapaata, A., Ndembi, N. and Rinke de Wit, T.F. (2012) Short communication: High prevalence of transmitted antiretroviral drug resistance among newly HIV type 1 diagnosed adults in Mombasa, Kenya. AIDS Research Human Retroviruses 28(9), 1033-1037.

Sluis-Cremer, N., Mark, A.W. and Raymond, F.S. (2015) Resistance to reverse transcriptase inhibitors used in the treatment and prevention of HIV-1 infection. Future Microbiol 10(11), 1773–1782.

Stanford University HIV drug resistance database (2021a) Major Integrase Inhibitor Resistance Mutations. Accessed at https://hivdb.stanford.edu/dr-summary/resistance-notes/INSTI/. Updated on 22/02/2021. HIVdb version 9.0.

Stanford University HIV drug resistance database (2021b) Major Nucleoside RT Inhibitor Resistance Mutations. Accessed at https://hivdb.stanford.edu/dr-summary/resistance-notes/NRTI/. Updated on 22/02/2021. HIVdb version 9.0

Stanford University HIV drug resistance database (2021c) Major Protease Inhibitor Resistance Mutations. Accessed at https://hivdb.stanford.edu/dr-summary/resistance-notes/PI/. Updated on 22/02/2021. HIVdb version 9.0

Stanford University HIV drug resistance database (2021) Major Non-Nucleoside RT Inhibitor Resistance Mutations, Accessed at https://hivdb.stanford.edu/dr-summary/resistance-notes/NNRTI/. Updated on 22/02/2021. HIVdb version 9.0

Steegen, K., Luchters, S., Dauwe, K., Reynaerts, J., Mandaliya, K., Jaoko, W., Plum, J., Temmerman, M. and Verhofstede, C. (2009) Effectiveness of antiretroviral therapy and development of drug resistance in HIV-1 infected patients in Mombasa, Kenya. AIDS Res Ther 6, 12.

Stephanie, W. and Alan, C. (2020) HIV: Guide to Protease Inhibitors. Healthline Media, Updated on April 25, 2020. Accessed at https://www.healthline.com/health/hiv-aids/protease-inhibitors.

Sulav, D., Christof, S. and Max v. K. (2012) Pharmacokinetics and pharmacodynamics of the reverse transcriptase inhibitor tenofovir and prophylactic efficacy against HIV-1 infection. PLoS One 7(7), e40382.

Susanna, L.L., Rebecca Rose, Sissy Cross, Christopher W. Rodriguez, Andrew D. Redd, Thomas C. Quinn, David Serwadda, Joseph Kagaayi, Godfrey Kigozi, Ronald Galiwango, Ronald H. Gray, M. Kate Grabowski and Laeyendecker., O. (2020) AIDS Research and Human Retroviruses 36(6), 522-526.

Tambuyzer, L., Nijs, S., Daems, B., Picchio, G. and Vingerhoets, J. (2011) Effect of mutations at position E138 in HIV-1 reverse transcriptase on phenotypic susceptibility and virologic response to etravirine. J Acquir Immune Defic Syndr 58(1), 18-22.

Tambuyzer, L., Vingerhoets, J., Azijn, H., Daems, B., Nijs, S., de Bethune, M.P. and Picchio, G. (2010) Characterization of genotypic and phenotypic changes in HIV-1-infected patients with virologic failure on an etravirine-containing regimen in the DUET-1 and DUET-2 clinical studies. AIDS Res Hum Retroviruses 26(11), 1197+.

Tamura, K. and Nei, M. (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution 10, 512-526.

Taylor, K., Fritz, K. and Parmar, M. (2021) Lamivudine. [Updated 2021 Jul 10]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls . Available from: https://www.ncbi.nlm.nih.gov/books/NBK559252/.

Tebit, D.M., Lobritz, M., Lalonde, M., Immonen, T., Singh, K., Sarafianos, S., Herchenroder, O., Krausslich, H.G. and Arts, E.J. (2010) Divergent evolution in reverse transcriptase (RT) of HIV-1 group O and M lineages: Impact on structure, fitness, and sensitivity to nonnucleoside RT inhibitors. J Virol 84(19), 9817–9830.

Thermo Fisher Scientific (2018) Applied Biosystems 3730xl DNA Analyzer with 3730xl Data Collection Software 5 User Guide.

ThermoFishers Scientific (2017) HIV-1 Genotyping Workflow Quick Reference, Pub. No. 100043578E. Catalog Numbers A32317 and A32318. Accessed at

https://tools.thermofisher.com/content/sfs/manuals/100043578_HIV1GenoWrk_QR.pdf.

Thompson, M.A., Aberg, J.A., Hoy, J.F., Telenti, A., Benson, C., Cahn, P., Eron, J.J., Jr., Günthard, H.F., Hammer, S.M., Reiss, P., Richman, D.D., Rizzardini, G., Thomas, D.L., Jacobsen, D.M. and Volberding, P.A. (2012) Antiretroviral treatment of adult HIV infection: 2012 recommendations of the International Antiviral Society–USA Panel. JAMA 308(4), 387–402.

Tiago, G., Gonzalo Bello, Paula Andrade, Ighor Arantes, João Marcos Pereira, Alexandre Bonfim Pinheiro da Silva, Rafael V. Veiga, Diana Mariani, Lídia Theodoro Boullosa, Mônica B. Arruda, José Carlos Couto Fernandez, Ann M. Dennis, David A. Rasmussen and Amilcar Tanuri (2021) HIV-1 molecular diversity in Brazil unveiled by 10 years of sampling by the national genotyping network. Sci Rep 11(15842).

Tsiang, M., Jones, G.S., Goldsmith, J., Mulato, A., Hansen, D., Kan, E., Luong, T., Rujuta, A.B., George, S., Kirsten, M.S., Niedziela-Majka, A., Stephen, R.Y., Helen, Y., , George, K., Tomas, C., Scott, E.L., Kirsten, L. W.,, and Haolun, J. (2016) Antiviral activity of bictegravir (GS-9883), a novel potent HIV-1 integrase strand transfer inhibitor with an improved resistance profile. Antimicrob Agents Chemother 60(12), 7086–7097.

United Nations Programme on HIV and AIDS (2018) AIDSinfo. Accessed August 2019.

United Nations Programme on HIV/AIDS (2017) Blind Spot, reaching out to men and boys. Addressing a blind spot in the response to HIV., Accessed at https://www.unaids.org/en/resources/campaigns/blind_spot.

Veronica, W., Michael Kiptoo, Joyceline Kinyua, Irene Odera, Edward Muge, Peter Muiruri, Raphael Lihana, Peter Kinyanjui and Elijah M Songok (2012) Predicted HIV-1 coreceptor usage among Kenya patients shows a high tendency for subtype d to be cxcr4 tropic. AIDS Research and Ther. 9(22).

Wainberg, M.A. (2004) The impact of the M184V substitution on drug resistance and viral fitness. Expert Review of Anti-Infective Therapy 2(1), 147-151.

World Health Organization (2006) Antiretroviral Therapy of HIV Infection in Infants and Children in Resource-Limited Settings, Towards Universal Access: Recommendations for a Public Health Approach. World Health Organization; Geneva, Switzerland: Feb, 2006

World Health Organization (2013a) Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection. Clinical guidance across the continuum of care: antiretroviral therapy. June 2013. Geneva, Switzerland

World Health Organization (2013b) First-line ART for pregnant and breastfeeding women and ARV drugs for their infants. Consolidated ARV guidelines.

World Health Organization (2018) Updated recommendations on first-line and second-line antiretroviral regimens and post-exposure prophylaxis and recommendations on early infant diagnosis of HIV.

World Health Organization (June 2013a) First-line ART for pregnant and breastfeeding women and ARV drugs for their infants. Consolidated ARV guidelines.

World Health Organization (June 2013b) Second-line ART for adults and adolescents. Consolidated ARV guidelines.

World Health organization and UNITAID (2017) Kenya to introduce better treatment for people living with HIV. Press release. 28 June 2017.

World Health Organization and United Nations Programme on HIV and AIDS (2019) The global health observatory data, summary of the global HIV epidemic. Accessed at: who.int/gho/hiv/en).

Yuncong, W., Hui Xing, Lingjie Liao, Zhe Wang, Bin Su, Quanbi Zhao, Yi Feng, Pengfei Ma, Jia Liu, Jianjun Wu, Yuhua Ruan and Yiming Shao (2014) The development of drug resistance mutations K103N Y181C and G190A in long term Nevirapine-containing antiviral therapy. volume 11: 36 (2014) AIDS Research and Therapy 11(36).

Zeh, C., Weidle, P.J., Nafisa, L., Lwamba, H.M., Okonji, J., Anyango, E., Bondo, P., Masaba, R., Fowler, M.G., Nkengasong, J.N., Thigpen, M.C. and Thomas, T. (2011) HIV-1 drug resistance emergence among breastfeeding infants born to HIV-infected mothers during a single-arm trial of triple-antiretroviral prophylaxis for prevention of mother-to-child transmission: a secondary analysis. PLoS Med 8(3), e1000430.

Zhang, M., Anne-Kathrin Schultz, Charles Calef, Carla Kuiken, Thomas Leitner, Bette Korber, Burkhard Morgenstern and Mario Stanke (2006) JPHMM at GOBICS: a web server to detect genomic recombinations in HIV-1 Nucleic Acids Research 34(suppl_2), W463–W465, https://doi.org/410.1093/nar/gkl1255.

Zoe, F., Dragsted, U.B., Gerstoft, J., Phillips, A.N., Kjaer, J., Mathiesen, L., Youle, M., Katlama, C., Hill, A., Bruun, J.N., Clumeck, N., Dellamonica, P. and Lundgren, J.D. (2006) A randomized trial to evaluate continuation versus discontinuation of

lamivudine in individuals failing a lamivudine-containing regimen: the COLATE trial. Antivir Ther 6(11), 761-770.

APPENDICES

Appendix 1: - Study Approval by the University

MASINDE MULIRO UNIVERSITY OF SCIENCE AND TECHNOLOGY
(MMUST)

Tel: 0702597360/61
: 0733120020/22
E-mail: deansgs@mmust.ac.ke
Website: www.mmust.ac.ke

P.O Box 190
50100 Kakamega
KENYA

Directorate of Postgraduate Studies

Ref: MMU/COR: 509079

5th February, 2018

Olipher Makwaga
SBB/H/03/15
P.O. Box 190-50100
KAKAMEGA

Dear Mr. Makwaga,

RE: APPROVAL OF PROPOSAL

I am pleased to inform you that the Directorate of Postgraduate Studies has considered and approved your Ph.D proposal entitled: ***"Prevalence and Molecular Characterization of HIV-1 Drug Resistance and Patterns Among Patients Visiting Selected Health Facilities in Busia County"*** and appointed the following as supervisors:

1. Dr. David Mulama - Department of Biological Sciences - MMUST
2. Dr. John Muoma - Department of Biological Sciences - MMUST
3. Prof. Matilu Mwau - KEMRI Busia

You are required to submit through your supervisor(s) progress reports every three months to the Director of Postgraduate Studies. Such reports should be copied to the following: Chairman, School of Natural Sciences & Technology Graduate Studies Committee and Chairman, Department of Biological Sciences. Kindly adhere to research ethics consideration in conducting research.

It is the policy and regulations of the University that you observe a deadline of three years from the date of registration to complete your Ph.D thesis. Do not hesitate to consult this office in case of any problem encountered in the course of your work.

We wish you the best in your research and hope the study will make original contribution to knowledge.

Yours Sincerely,

Prof. John Obiri
DIRECTOR DIRECTORATE OF POSTGRADUATE STUDIES

Appendix 2:- Study approval by SERU

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840-00200, NAIROBI, Kenya
Tel: (254) (020) 2722541, 2713349, 0722-205901, 0733-400003, Fax: (254) (020) 2720030
E-mail: director@kemri.org, info@kemri.org, Website. www.kemri.org

KEMRI/RES/7/3/1 **December 14, 2016**

TO: **OLIPHER MAKWAGA,**
PRINCIPAL INVESTIGATOR

THROUGH: **ACTING DIRECTOR, CIPDCR,**
BUSIA

Dear Madam,

RE: **PROTOCOL NO. KEMRI/SERU/CIPDCR/008/3333 (*RESUBMISSION2 OF INITIAL SUBMISSION*): PREVALENCE AND MOLECULAR CHARACTERIZATION OF HIV-1 DRUG RESISTANCE AMONG PATIENTS VISITING SELECTED RURAL HEALTH FACILITIES IN BUSIA COUNTY, KENYA _ (*VERSION 4.0 DATED 3RD NOVEMBER, 2016*)**

Reference is made to your letter dated 6th November, 2016. The KEMRI/Scientific and Ethics Review Unit (SERU) acknowledges receipt of the revised study documents on the 8th November, 2016.

This is to inform you that the Committee notes that the issues raised during the 254th Committee C meeting of the KEMRI/SERU held on 21st August, 2016 have been adequately addressed.

Consequently, the study is granted approval for implementation effective this day, **14th December, 2016** for a period of one year. Please note that authorization to conduct this study will automatically expire on **December 13, 2017.** If you plan to continue data collection or analysis beyond this date, please submit an application for continuation approval to SERU by **30th October, 2017**.

You are required to submit any proposed changes to this study to SERU for review and the changes should not be initiated until written approval from SERU is received. Please note that any unanticipated problems resulting from the implementation of this study should be brought to the attention of SERU and you should advise SERU when the study is completed or discontinued.

You may embark on the study.

Yours faithfully,

for

DR. EVANS AMUKOYE,
ACTING HEAD,
KEMRI/SCIENTIFIC AND ETHICS REVIEW UNIT

Appendix 3: - Study Renewal of Approval

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840-00200, NAIROBI, Kenya
Tel: (254) 2722541, 2713349, 0722-205901,0733-400003, Fax: (254) (020) 2720030
Email: director@kemri.org, info@kemri.org, Website. www.kemri.org

KEMRI/RES/7/3/1 **July 01, 2020**

TO: **OLIPHER MAKWAGA,**
PRINCIPAL INVESTIGATOR.

THROUGH: **THE DEPUTY DIRECTOR, CIPDCR,**
BUSIA.

Dear Madam,

RE: **KEMRI/SERU/CIPDCR/008/3333 *(REQUEST FOR ANNUAL RENEWAL)*: PREVALENCE AND MOLECULAR CHARACTERIZATION OF HIV-1 DRUG RESISTANCE PATTERNS AMONG PATIENTS VISITING SELECTED RURAL HEALTH FACILITIES IN BUSIA COUNTY, KENYA.**

Thank you for the continuing review report for the period **May 20, 2019** to **May 19, 2020.**

This is to inform you that the Expedited Review Team of the KEMRI Scientific and Ethics Review Unit (SERU) was of the informed opinion that the progress made during the reported period is satisfactory. The study has therefore been granted **approval**.

This approval is valid from **July 01, 2020** through to **June 30, 2021**. Please note that authorization to conduct this study will automatically expire on **June 30, 2021**. If you plan to continue with data collection or analysis beyond this date, please submit an application for continuing approval to SERU by **May 19, 2021.**

You are required to submit any amendments to this protocol and other information pertinent to human participation in this study to SERU for review prior to initiation.

You may continue with the study.

Yours faithfully,

ENOCK KEBENEI,
THE ACTING HEAD,
KEMRI SCIENTIFIC AND ETHICS REVIEW UNIT.

Appendix 4: - DNA Sequences of Controls through Cycle Sequencing

Sequence of RNA positive control

Note: The following sequence is of the full synthesized RNA control. The HIV-1 *pol* fragment that is sequenced using this genotyping kit is bolded with the starting and ending codons underlined.

The sequence encompasses:

- Protease codons 6 to 99
- Reverse Transcriptase codons 1 to 251

```
   1 ACCAAATGAA AGATTGTACT GAGAGACAGG CTAATTTTTT AGGGAAGATC TGGCCTTCCT ACAAGGGAAG
  71 GCCAGGGAAT TTTCTTCAGA GCAGACCAGA GCCAACAGCC CCACCAGAAG AGAGCTTCAG GTCTGGGGTA
 141 GAGACAACAA CTCCCCCTCA GAAGCAGGAG CCGATAGACA AGGAACTGTA TCCTTTAACT TCCCTCAGGT
 211 CACTCTTTGG CAACGACCCC TCGTCACAAT AAAGATAGGG GGGCAACTAA AGGAAGCTCT ATTAGATACA
 281 GGAGCAGATG ATACAGTATT AGAAGAAATG AGTTTGCCAG GATGATGGAA ACCAAAAATG ATAGGGGGAA
 351 TTGGAGGTTT TATGAAAGTA AGACAGTATG ATCAGATACT CATAGAAATC TGTGGACATA AAGCTATAGG
 421 TACAGTATTA GTAGGACCTA CACCTGTCAA CATAATTGGA AGAAATCTGA TGACTCAGAT TGGTTGCACT
 491 TTAAATTTTC CCATTAGCCC TATTGAGACT GTACCAGTAA AATTAAAGCC AGGAATGGAT GGCCCAAAAG
 561 TTAAACAATG GCCATTGACA GAAGAAAAAA TAAAAGCATT AGTAGAAATT TGTACAGAGC TGGAAAAGGA
 631 AGGGAAAATT TCAAAAATTG GGCCTGAAAA TCCATACAAT ACTCCAGTAT TTGCCATAAA GAGAAAAGAC
 701 AGTACTAAAT GGAGAAAATT AGTAGATTTC AGAGAACTTA ATAAGAGAAC TCAAGACTTC TGGGAAGTTC
 771 AATTAGGAAT ACCACATCCC GCAGGGTTAA AAAAGAATAA ATCAGTAACA GTACTGGATG TGGGTGATGC
 841 ATATTTTTCA GTTCCCTTAG ATGAAGACTT CAGGAAGTAT ACTGCATTTA CCATACCTAG TATAAACAAT
 911 GAGACACCAG GGATTAGATA TCAGTACAAT GTGCTTCCAC AGGGATGGAA AGGATCACCA GCAATATTCC
 981 AAAGTAGCAT GACAAAAATC TTAGAGCCTT TTAGAAAACA AAATCCAGAC ATAGTTATCT GTCAATACGT
1051 GGATGATTTG TATGTAGGAT CTGACTTAGA AATAGGGCAG CATAGAACAA AAATAGAGGA GCTGAGACAA
1121 CATCTGTTGA GGTGGGGACT TACCACACCA GACAAAAAAC ATCAGAAAGA ACCTCCATTC CTTTGGATGG
1191 GTTATGAACT CCATCCTGAT AAATGGACAG TACAGCCTAT AGTGCTGCCA GAAAAAGACA GCTGGACTGT
1261 CAATGACATA CAGAAGTTAG TGGGGAAATT GAATTGGGCA AGTCAGATTT ACCCAGGGAT TAAAGTAAGG
1331
```

Mutation profile of RNA positive control

The positive control contains the following mutations (mutations in bold cause varying levels of drug resistance):

- Protease: N37S, R41*, **154M**, and **L90M**

  Note: * Protease position 41 **should** contain a stop codon.
- Reverse transcriptase: **M41L**, **K65R**, **K103N**, K122E, **Y181C**, **M184V**, and F214L

Partial sequence of pGEM DNA sequencing control

```
  1 GAATTGTAAT ACGACTCACT ATAGGGCGAA TTCGAGCTCG GTACCCGGGG ATCCTCTAGA GTCGACCTGC
 71 AGGCATGCAA GCTTGAGTAT TCTATAGTGT CACCTAAATA GCTTGGCGTA ATCATGGTCA TAGCTGTTTC
141 CTGTGTGAAA TTGTTATCCG CTCACAATTC CACACAACAT ACGAGCCGGA AGCATAAAGT GTAAAGCCTG
211 GGGTGCCTAA TGAGTGAGCT AACTCACATT AATTGCGTTG CGCTCACTGC CCGCTTTCCA GTCGGGAAAC
281 CTGTCGTGCC AGCTGCATTA ATGAATCGGC CAACGCGCGG GGAGAGGCGG TTTGCGTATT GGGCGCTCTT
351 CCGCTTCCTC GCTCACTGAC TCGCTGCGCT CGGTCGTTCG GCTGCGGCGA GCGGTATCAG CTCACTCAAA
421 GGCGGTAATA CGGTTATCCA CAGAATCAGG GGATAACGCA GGAAAGAACA TGTGAGCAAA AGGCCAGCAA
491 AAGGCCAGGA ACCGTAAAAA GGCCGCGTTG CTGGCGTTTT TCCATAGGCT CCGCCCCCCT GACGAGCATC
561 ACAAAAATCG ACGCTCAAGT CAGAGGTGGC GAAACCCGAC AGGACTATAA AGATACCAGG CGTTTCCCCC
631 TGGAAGCTCC CTCGTGCGCT CTCCTGTTCC GACCCTGCCG CTTACCGGAT ACCTGTCCGC CTTTCTCCCT
701 TCGGGAAGCG TGGCGCTTTC TCATAGCTCA CGCTGTAGGT ATCTCAGTTC GGTGTAGGTC GTTCGCTCCA
771 AGCTGGGCTG TGTGCACGAA CCCCCCGTTC AGCCCGACCG CTGCGCCTTA TCCGGTAACT ATCGTCTTGA
841 GTCCAACCCG GTAAGACACG ACTTATCGCC ACTGGCAGCA GCCACTGGTA ACAGGATTAG CAGAGCGAGG
911 TATGTAGGCG GTGCTACAGA GTTCTTGAAG TGGTGGCCTA ACTACGGCTA CACTAGAAGG ACAGTATTTG
981 GTATCTGCGC TCTGCTGAAG
```

Appendix 5: - Base calling using Recall software


Sample 141 (id: 515066)
Close map
Standard
R3_2020-07-14
R2_2020-07-14
F1_2020-07-14
F3_2020-07-14
F2_2020-07-14
R1_2020-07-14
(243) Y Y
(244) Q Q
(245) Y Y
(246) N N
(247) V V
(248) L L
(249) P P
(250) R_Q151 Q
(251) G G
Reference Protein
A T A T C A G T A C A A T G T G C T T C C A C A G G G A
G T A T C A G T A C A A A T G T G C T T C C R C A G G G A
Standard
Assembled
F3_2020-07-14 (reverse)
F2_2020-07-14 (forward)
R1_2020-07-14 (forward)
722
Job name: Aug-11-THERMO_HIV (id: 115130)
Upload date: 2021-08-11
Status: Passed
Mixtures: 12 (cutoff: 17.5%)
Marks: 4
"N"s: 0
Edited bases: 0
Errors: 0
Use the following keys to navigate:
Next marked base: right arrow
Previous marked base: left arrow
Next base: shift + right arrow
Previous base: shift + left arrow
With key locations defined in advanced settings:
Next marked key base: down arrow
Previous marked key base: up arrow
Use the following keys to make edits:
Change base: A,C,G,T,N R,Y,K,M,S,W,B,D,H,V
Erase base: dash
Mixture compositions:
R = A/G Y = C/T K = G/T M = A/C
S = G/C W = A/T B = C/G/T D = A/G/T
H = A/C/T V = A/C/G N = A/C/G/T
Jump to base
Jump to AA
Save & Pass
Fail Sample
Exit


Sample A82 (id: 515076)
Close map
Standard
R3_2020-07-14
F1_2020-07-14
F3_2020-07-14
R2_2020-07-14
F2_2020-07-14
R1_2020-07-14
(123) W W
(124) P P
(125) L L
(126) T T
(127) E E
(128) E E
(129) K K
(130) I I
(131) K K
Reference Protein
A T G G C C A T T G A C A G A A G A A A A A A T A A A A
A T G G C C R T T G A C A G A A G A A A A A A T A A A A
Standard
Assembled
R3_2020-07-14 (reverse)
F1_2020-07-14 (forward)
F3_2020-07-14 (reverse)
F2_2020-07-14 (forward)
362
Job name: Aug-11-THERMO_HIV (id: 115130)
Upload date: 2021-08-11
Status: Passed
Mixtures: 14 (cutoff: 17.5%)
Marks: 10
"N"s: 0
Edited bases: 0
Errors: 0
Use the following keys to navigate:
Next marked base: right arrow
Previous marked base: left arrow
Next base: shift + right arrow
Previous base: shift + left arrow
With key locations defined in advanced settings:
Next marked key base: down arrow
Previous marked key base: up arrow
Use the following keys to make edits:
Change base: A,C,G,T,N R,Y,K,M,S,W,B,D,H,V
Erase base: dash
Mixture compositions:
R = A/G Y = C/T K = G/T M = A/C
S = G/C W = A/T B = C/G/T D = A/G/T
H = A/C/T V = A/C/G N = A/C/G/T
Jump to base
Jump to AA
Save & Pass
Fail Sample
Exit

Appendix 6: - Sequences of HIV-1 DNA in FASTA Form

Sample_1 > Busia 1_A_1 _ D

TGGCAACGACCAGTTGTCACAGTAAAAATAGAGGGACAACTAAAAGAAGCTTTATTAGACAC
AAGAGCAGATGATACAGTACTAGAAGATATAAATTTACCAGGAAAATGGAAACCAAGAATGA
TAGGGGAAAGTGGAGGCTTTATCAAGGTAAGGCAATATGATCAAATAGTGATAGAAATTTGT
GGGAAAAAGGCTATAGGTACAGTGTTAATAGGACCTACACCTGTCAACATAATTGGCAGAAA
TATGTTGACTCAGATTGGTTGTACTTTAAATTTTCCTATTAGTCCTATTGAGACTGTACCAGTA
AAATTAAAACCAGGAATGGATGGCCCAAAAGTTAAACAATGGCCATTAACAGAAGAAAAAAT
AAAAGCATTAACAGAAATTTGTATGGAAATGGAAAAGGAAGGAAAAATCTCAAAAATTGGGC
CTGAAAATCCATACAATACTCCAGTATTTGCTATAAAGAAAAAAGACAGTACTAAATGGAGG
AAATTAGTAGATTTCAGAGAGCTCAATAAAAGAACTCAGGACTTCTGGGAAGTTCAATTAGG
AATACCTCATCCAGCAGGCTTAAAAAAGAATAAATCAGTAACAGTACTAGATGTGGGGGATG
CATATTTTTCAGTGCCTTTAGATGAGAACTTTAGAAAGTATACTGCATTCACCATACCTAGTAT
AAACAATGAGACACCAGGAATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGGT
CACCAGCAGTATTCCAAAGTAGCATGACAAAAATCTTAGAACCTTTTAGAAAACAAAACCCA
GAAATAGTCATCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAG
CATAGAACAAAAGTAGAAGAATTGAGGGAACACCTCTTGAAGTGGGGATTTACCACACCAGA
CAAAAAGCATCAGAAAGAACCTCCATTTCTTTGGATGGGTTATGAACTCCATCCTGATAAATG
G

Sample_2 > Busia 2_A_1 _ K

GTGATTGTTTTTCCTGGACTAAAAAAATGGGGGAAGTTAAAGAAGTTTTTTTAGGACCAGACC
AGTTGTTACATTTTAGAGAATTAATTTCCAGGGAAAGGGAAACAAAAATGATAGGGGGAATT
GGGGTTTTCATCAATGTAAAACAGTATGATCAAATGGTAATGGAAATTTGTGGAAAAAAGGCT
ATAGGGACAGTATTAGTAGGACCTACACCTGTCGGCATAATTGGAAGAAAACTGTTGACCCA
GATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTACCAGTACAATGAAACCCA
GGAATGGATGGCCCAACGGCGAAACAAAAAACAGTGACAGAAGAAAAAATAAAAGCGTTAA
CAGAAATTTGTTTAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCCTGAAAATCCA
TACTATACTCCAATATTTGCAATAAAGAAAAAATATAGCACTAATTGGAGTATATTAGTAGAT
TTCAGATATCTCTTTAACAGAACACTAGACTTTTGGGAAGTTCAATTAGGAATACCGCATCCA
GCGGGCCTACAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGACGCATATTTTTCATTT
CCTCTACATGAAAGCTTTAGAAAGTATACTGCATTCACCATACCTAGTACAAACAATGAGACA
CCATGAATTATGTATCATTACAATGTGCTTCCACGGGGATGGAAAGGATCACCAGCAATATTC
CAGAGTAGCATGACAAAAATCTTAGAGCCATTTAGATCAAAAAACCCAGAAATAGTTATCTAT
CAATACATGGATGATTTGTATGTAGGATCTGATTTAGAAATAGGGCAGCATAGAACAAAAAT
AGAAGAGTTGAGAGCTCATCTATTGAGCTGGAGATTAACTACACCAGACAAAAAGCATCAGA
AAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG

Sample 3 >Busia3_A1

TGGCAACGACCCCTTGTCACAGTAAAAATAGGAGGACAGTTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAAATTTACCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAAATACTTATAGAAATTTGTG
GAAATAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTGTTAAAACTGTACCAGTAA
AATTAAAACCAGGAATGGATGGCCCAAAGGTTAAGCAATGGCCATTGACAGAAGAGAAAATA
AAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAGATCTCAAAAATTGGGCC
TGAAAATCCATACAATTCTCCAATATTTGCTATAAAGAAGAAAAATAGCAATAGATGGAGAA
AATTAGTAGATTTTAGAGAACTCAACAAAAGAACTCAGGACTTCTGGGAAGTTCAATTAGGA
ATACCTCATCCAGCAGGTTTAAAGAAGAAAAAATCAGCAACAGTAATAGATGTGGGGGATGC

ATATTTTTCAGTACCTTTAGATGAAGACTTTAGAAAGTATACTGCATTCACCATACCTAGTACA
AACAATGAGACACCAGGAGTCAGGTATCAATACAATGTGCTGCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCTTTCAGATCACAAAATCCAGA
CATAGTTATCTATCAATACGTGGATGACTTATATGTAGGATCTGATTTAGAAATAGGGCAGCA
TAGAACAAAAGTAAAAGAGTTAAGAGCTCATCTATTGAGCTGGGGGTTTTTTACACCAGACCA
AAAGCATCAGAAAGAACCTCCTTTGCTATGGATGGGATATGAACTGCATCCTGACAAATGG
Sample 4>Busia4_D
TGGCAACGACCCCTCGTCACAGTAAAAATAGAGGGACAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGCTTTATCAAAGTAAGACAGTATGATCAGGTACATGTAGACATCTGTG
GACATAAAGCTGTAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATT
TGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTCCTATTGAAACTGTACCAGTAAA
ATTAAAGCCAGGGATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAATAA
TAGCACTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAGAATTGGGCCT
GAAAATCCATATAATACTCCAATATTTGCCATAAAGAGAAAAGATGGTATTAAGTGGAGAAA
ATTAGTAGATTTCAGAGAACTCAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAAT
ACCACATCCTGCAGGGCTAAAAAAGAAAAAATCAGTAACAATACTGGATGTGGGTGATGCAT
ATTTTTCAGTTCCCTTAGATGAAGACTTCAGGAAGTATACTGCATTTACCATACCTAGTAGAAA
CAATGAGACACCAGGAATTAGATATCAGTACAATGTGCTTCCACAAGGATGGAAAGGGTCAC
CAGCAATATTTCAAAGTAGCATGACAAAAATCTTAGAGCCTTTTAGAAAACAAAATCCAGAC
ATGATTATATGTCAATACATGGATGATTTGTATGTAGGACCTGACTTAGAAATAGGACAGCAT
AGAAGAAAAATAGAGGAGCTAAGGGAACACCTATTGAAGTGGGGAATTACCACACCAGACA
AGAAACATCAGAAAGAACCTCCATTTCTTTGGATGGGCTATGAACTCCATCCTGATAAATGG
Sample 5 >Busia5_D
TGGCAACGACCCCTTGTCACAATAAAGATAGGGGGACAGATAAGGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAGATTTGCCAGGAAAGTGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAAATACTTGTAGACATATGTG
GACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATT
TGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTTCTATTGAAACTGTACCAGTAAA
ATTAAAGCCAGGAATGGATGGCCCGAAAGTAAAACAATGGCCATTGACAGAGGAAAAAATA
AAAGCACTAATAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCCATAAAGAAAAAGGACAGTACTAAGTGGAGAA
AATTAGTAGATTTCAGGGAACTTAATAAGAGAACTCAAGATTTCTGGGAAGTTCAACTAGGAA
TACCACATCCTGCAGGGCTAAAAAAGAACAAGTCAGTAACAGTACTGGATGTGGGTGATGCA
TATTTTTCAGTTCCCTTGGATAAAGACTTTAGAAAATATACTGCATTCACCATACCCAGTATAA
ACAATGAGAAACCAGGAACTAGATATCAGTACAATGTGCTTCCACAAGGATGGAAAGGATCA
CCAGCAATATTCCAATGTAGCATGACAAAAATCTTAGAACCTTTTAGACAACAGAATCCAGAC
ATAGTTATTTATCAATACGTGGATGATTTATATGTAGGATCTGATTTAGAAATAGGACAGCAT
AGAACAAAAATAGAGGAATTAAGAGAACACTTATTGAAGTGGGGATTTACCACACCAGACAA
AAAGCATCAGAAAGAACCCCCATTTCTTTGGCTGGGTTATGAACTCCATCCTGATAAATGG
Sample 6 >112_A1
TGGCAACGACCCCTTGTCACAGTAAGGATAGAAGGACAACTAAAAGAGGCTCTCTTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAAACTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATTAAGGTAAAACAGTATGATCAGATAGCTATTGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGGAAT
ATTTTGACTCAAATTGGTTGTACTTTAAATTTTCCAATAAGTCCTATTGAGACTGTACCAGTAA
CATTAAAGCCAGGAATGGATGGCCCTAGGATTAAACAATGGCCATTGTCAGAAGAAAAAATA
AAAGCATTGACAGAAATTTGTGCAGATATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC

TGAAAATCCATACAATACTCCAATATTTGCAATAAAGAAAAAAGATAGCACTAAATGGAGAA
AATTAGTAGATTTCAGAGAGCTTAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGGA
ATACCGCATCCAGCGGGCCTAAAAAAGAACAAATCAGTAACAGTACTGGATGTGGGGGATGC
ATATTTCTCAGTTCCTTTAGATGAAAATTTTAGAAAATATACTGCATTCACCATACCTAGTACA
AACAATGAGACACCAGGAATCAGGTTTCAGTACAATGTGCTTCCACAAGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAATCCAG
AAATAGTTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGGAGC
ACAGAACAAAAATAGAAGAATTGAGAGCTCATCTATTGAGCTGGGGATTTACTACCCCAGAC
AAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG

Sample 7 >Busia7_A2D

TGGCAACGACCCCTTGTACCAAAAAAAAAAAGAGGGGGGGGAGGGGCCCCCCGGGTTTTTAG
GGTTGTGGCGGGGGTCGTATTAGAAGATATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATTAAAGTAAGACAGTATGATCAAGTACCCGTAGAAATCTGTG
GATATAAAGCTGTAGGTACAGTATTGGTAGGACCTACACCTGTCAACATAATTGGAAGAAATT
TGTTGACTCAAATTGGTTGTACTTTAAATTTTCCAATCAGTCCTATTGAAACTGTACCAGTAAA
ATTAAAGCCAGGGATGGATGGCCCAAAAGTTAAACAATGGCCGTTGACAGAAGAAAAAATAA
AAGCACTAACAGAAATTTGTACAGAAATGGAAAAGGAGGGAAAAATTTCAAGAATTGGGCCT
GAAAATCCATACAATACTCCAATATTTGCCATAAAGAAAAAAGACAGTACCAAGTGGAGGAA
ATTAGTAGATTTCAGAGAACTTAACAAGAGAACTCAAGACTTTTGGGAAGTTCAACTAGGAAT
ACCACATCCTGCAGGGCTAAAAAAGAACAAATCAGTAACAGTACTGGATGTGGGTGATGCAT
ATTTTTCAGTTCCCTTATATGAAGACTTTAGAAAATATACCGCATTCACCATACCTAGTGTAAA
CAATGAGACACCAGGAATTAGATATCAATACAATGTGCTTCCACAGGGATGGAAAGGATCAC
CGGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAACCTTTTAGAAAACAAAATCCAGGA
ATAGTTATCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATTGGGCAGCAT
AGAATAAAAATAGAAGAGTTAAGGGGACACCTCTTAAAGTGGGGATTTACCACACCAGACAA
AAAACATCAGAAAGAACCTCCATTTCTTTGGATGGGTTATGAACTCCATCCTGATAAATGG

Sample 8>Busia8_A1B

TGGCAACGACCCCTCGTCACAATAAAGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATGAGTTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAGGTAAGACAATATGATCAGGTACTTCTAGAAATTTGTG
GAAAAAAGGCTATAGGTTCAGTATTAGTAGGACCTACACCTGTTAACATAATTGGAAGAAAT
ATGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTTCTATTGAGACTGTGCCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGATAAAACAATGGCCATTGACAGAAGAAAAAAT
AAAAGCATTAACAGAAATTTGTAAAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGC
CTGAAAATCCATATAATACTCCAGTCTTTGCTATAAAGAAAAAAGATAGCACTAAATGGAGA
AAGTTAGTAGATTTTAGAGAGCTCAATAAAAGAACTCAGGACTTCTGGGAAGTCCAATTAGG
AATACCGCATCCCGCAGGTTTAAAAAAGAACAAATCAGTAACAGTACTAGATGTTGGGGATG
CATATTTTCAGTTCCCTTAGATGAAAACTTTAGAAAGTATACTGCATTCACCATACCCAGTATA
AACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCGGCAATATTCCAGAGTAGCATGACAAAAATCTTGGAGCCCTTTAGATCACAAAATCCAGA
AATAATTATCTATCAATACATGGATGATTTGTATGTAGGATCTGATTTAGAAATAGGGCAGCA
TAGAACAAAAGTAGAGGAGTTGAGAGCTCACCTATTGAGCTGGGGATTTACTACACCAGACA
AAAAGCATCAGAAAGAACCTCCATTTCTTTGGATGGGATATGAACTCCATCCTGACAAATGG

Sample 9 >Busia9_A1

TGGCAACGACTTCTGGTCACAGTAAAAATAGGGGGACAGCTAAAAGAAGCTCTATTAGACAC
AGGAGCAGATGATACAGTACTAGAAGACATAAGTTTGCCAGGGAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAGGTAAAACAGTATGATCAGATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTGGTAGGACCTACACCTGTCAACATAATTGGAAGGAAT

ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAAACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCTAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTAAAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGATAGCACAAAATGGAGAA
AGTTAGTAGATTTCAGAGAACTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGAA
TACCGCATCCAGCGGGCTTAAAAAGGAACAAATCAGTAACAGTATTAGATGTGGGGGATGCG
TACTTTTCAGTTCCCCTAGATAAAGATTTTAGAAAGTATACTGCATTCACCATACCTAGCATAA
ACAATGAGACACCAGGAATCAGGTATCAGTATAATGTGCTTCCACAAGGATGGAAAGGATCA
CCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAGCAAAAAATCCAGA
CATAGTTATCTATCAATACGTGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGCA
TAGAACAAAAATAGAAGAGTTAAGAGAACATCTCTTGAGATGGGGGTTTACTACACCAGACA
AAAAGCATCAGAAAGAACCTCCATTTCTTTGGATGGGATATGAACTCCATCCGGATAAGTGG
Sample 10 >Busia10_A1
TGGCAACGACCCCTCGTCACAATAAAAATAGGGGGACAGCTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTCATCAAAGTAAAACAGTATGATCAGGTACTTCTAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTGGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACCCAGATTGGTTGTACTTTAAATTTTCCCATAAGTCCTATTCCTACTGTACCAGTAA
GTTTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAATAGAAATTTGTACAGATATGGAAAAGGAAGGAAAAATTTCAAGAGTTGGGCC
TGAAAACCCATATAATACTCCAATATTTGCTATAAAGAAAAAAGACAGCACTAAATGGAGAA
AATTAGTAGATTTCAGAGAGCTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGAA
TACCGCATCCAGCGGGTCTAAAGAAGAACAAATCAGTAACAGTACTAGATGTGGGGGATGCA
TATTTTTCAGTTCCCTTAGATGAAAACTTTAGAAAGTATACAGCATTCACTATACCTAGTACAA
ACAATGAAACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCC
CCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAATCCAGA
AATAATTATCTATCAATACATGGATGATTTATATGTGGGATCTGACTTAGAAATAGGACAGCA
TAGAACAAAAGTAGAGGAGCTAAGAGAGCATCTATTGAGATGGGGGTTTACCACACCAGATA
AAAAGCATCAGAAAGAACCTCCATTTCTTTGGATGGGATATGAACTCCATCCTGACAAGTGG
Sample 11 >Busia11_A1
TGGCAACGACCCCTCGTCACAATAAAGATAGGGGGACAGATAGTAGAAGCTCTATTAGACAC
AGGAGCAGATGATACAGTATTAGAAAACATAAATTTGCCAGGAAAATGGAAACCAAAAATAA
TAGGGGGACTTGGAGGTTTAATCAGAGTAAAAGAGTATGATCAGATACCTATAGAATTTTGTG
GGAAAAAGGTTATAGGTTCAGTATTAGTAGGACCTACACCTGCCAACATAATTGGGAGAAAT
ATTTTGACCCAGATTGGTTGTACTCTAAATTTCCCAATTAGTCATATTGAGACTGTACCAGTGA
AATTAAAGCCAGGAATGGATGGCCCAAGAGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATCTGTACAGAATTGGAAAAGGAAGGGAAAATTTCAAAAATTGGGCC
TGAGAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAGAACAGCACTAGATGGAGAA
AACTAGTAGATTTCAGAGAACTCAATAAAAGAACTCAAGACTTTTGGGAAGTGCAATTAGGA
ATACCGCATCCAGCGGGCTTAAAAAAGAACAAATCAGTAACAGTACTAGATGTGGGGGACGC
ATATTTTTCAATTCCCTTAGATGAAAACTTTAGAAAGTATACAGCATTCACCATACCTAGTACA
AACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGGTC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAACCCAG
GAATAACTATCTATCAATACGTGGATGACTTGTACGTAGGGTCTGATTTAGAAATAGGACAGC
ATAGAGCAAAAATAGAAGAGCTAAGAGCTCATCTATTGAGCTGGGGACTGTATACACCAGAC
CAAAAGCATCAGAAAGAGCCTCCATTTCATTGGATGGGATATGAACTCCATCCTGACAAGTGG
Sample 12 >Busia12_A1

TGGCAACGACCCCTTGTTGCCGTGAGAATAGGGGGACAGGTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTCGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGGTTTATCAAGGTAAAACAGTATGATCAGGTACTTATAGACATTTGTG
GGAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTAAACATAATTGGAAGGAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTACCAGTAA
CATTAAAGCCAGGAATGGATGGCCCAAGAATTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTAAAGAGATGGAAAAGGAAGGAAAAATTTCAAAGATTGGGCC
TGAAAATCCATATAATACCCCAATATTTGCTATAAAGAAAAAAGATAGCACTAAATGGAGAA
AGTTAGTAGATTTTAGAGAACTCAATAAAAGAACTCAGGACTTCTGGGAAGTTCAATTAGGAA
TACCACATCCCGCAGGTTTAAAAAAGAAAAAATCAGTAACTGTACTAGATGTGGGGGACGCA
TACTTTTCAGTCCCTTTAGATGAAAGCTTTAGAAAGTACACTGCATTCACCATACCTAGTACAA
ACAATGAGACACCAGGAATCAGGTATCAGTATAATGTGCTTCCACAGGGGTGGAAAGGATCA
CCTGCGATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTCAGGTCAAAAAATCCAGAA
ATAGTTATCTGTCAATACGTGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGCAT
AGAGCAAAAGTAGAAGAATTAAGACAACATCTGTTGAGTTGGGGGTTTACTACACCAGACAA
AAAGCATCAGAAAAAACCTCCATTCCTTTGGATGGGTTATGAACTCCATCCTGATAAATGG
Sample 13 >Busia13_A1B
TGGCAACGACCCCTCGTCACAATAAAGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATAC
AGGAGCTGATGATACAGTATTAGAAGATATAGATTTGCCAGGAAGATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACCTATAGAAATTTGTG
GGAAAAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGATTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTAATGATATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCAATAAAAAAGAAAGATAGCACTAAATGGAGAA
AATTAGTAGATTTCAGAGAACTCAATAAAAGAACACAAGACTTTTGGGAAGTTCAACTAGGA
ATACCACATCCAGCGGGCCTAAAAAAGAAAAAGTCAGTAACAGTACTAGATGTGGGGGATGC
ATATTTTTCAGTTCCTTTACATGAAGACTTTAGAAAATATACTGCATTCACCATACCTAGTATA
AACAATGAGACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAAAAAAAAATCCAG
AAATAGTCATCTATCAATACATGGATGACTTGTATGTAGGATCAGATTTAGAAATAGGGCAGC
ATAGAACAAAAATAGAGGAGCTGAGACAACATCTGTTGAGGTGGGGACTTACCACACCAGAC
AAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGAACTCCATCCTGATAAATGG
Sample 14 >Busia14_D
TGGCAACGACCCCTCGTCACAATAAAGATAGGGGGGACACTGAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAAACCTGCCAGGAAAATGGAAACCAAAAATG
ATAGGGGGAATTGGAGGCTTTATCAAAGTAAGACAGTATGATAATATACAAGTAGACATTTAT
GGACATAAAGCTGTAGGTACAGTGTTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAA
TTTGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTCCTATTGAAACTGTACCAGTA
AAATTAAAGCCAGGGATGGATGGCCCAAAAGTTAAACAATGGCCACTGACAGAAGAGAAAAT
AAAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATYTCAAGAATTGGG
CCTGAAAATCCATACAATACTCCAATATTTGCCATAAAGAAAAAGGATAGTACTAAATGGAG
AAAATTAGTAGATTTTAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGG
AATACCACATCCTGCAGGGCTAGAAAAGAACAAGTCAATAACAGTAATAGATGTGGGTGATG
CATATTTTTCAGTTCCCTTAGATGAAAACTTTAGAAAATATACAGCATTCACCATACCTAGTAC
AAACAATGGGACACCAGGAATTAGATATCAGTACAATGTGCTCCCACAAGGATGGAAAGGAT
CACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAACCTTTTAGAAAACAAAATCCA
GAGATGGTTATCTATCAATACGTGGATGATTTATATGTAGCATCTGACTTAGAAATAGGGCAG

CATAGAAAGAAAATAGAGGAATTAAGGGAACACCTATTAAAGTGGGGATTTACCACGCCAGA
CAAGAAACATCAGAAAGAACCTCCCTTTCTTTGGATGGGTTATGAACTCCATCCTGATAAATG
G
Sample 15 >Busia15_A1
TGGCAACGACCCATTGTTACAATAAAAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAGAGTTGCCAGGAAAATGGAAACCAAGAATG
ATAGGGGGAATTGGAGGTTTCATCAAAGTGAGACAATATGATCAGGTAGTTATAGAAATTTGT
GGGAAAAAGGCCATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAA
CATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTGCCAGTA
AAATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAAT
AAAAGCATTAACAGAAATTTGTACAGAGATGGAAAGGGAAGGAAAAATTTCAAGAATTGGGC
CTGAAAATCCATACAATACGCCAATATTTGCAATAAAGAAAAAAGATAGCACTAAATGGAGG
AAATTAGTAGATTTCAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAAGTTCAGTTAGG
GATACCGCATCCAGCGGGTCTAAAAAAGAACAAATCAGTAACAGTATTAGATGTGGGGGATG
CATATTTTTCAGTTCCTTTAGATGAAAGCTTTAGAAAATATACTGCATTCACCATACCTAGTAC
AAACAATGAGACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGAT
CACCAGCAATATTCCAGAGTAGTATGACAAAAATCTTAGAGCCATTTAGATTAAAGAATCCAG
AAATAATTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGACAAC
ATAGAACAAAAGTAGAAGAGTTGAGAGGTCATCTATTGAGCTGGGGGTTTACTACACCAGAT
AAAAAGCATCAGAAAGAACCTCCATTCCGTTGGATGGGGTATGAGCTCCATCCTGACAAGTG
G
Sample 16 >Busia16_A1
TGGCAACGACCCCTTGTCACAGTAAAAATAGGAGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAAATTTGCAGGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAGGTAAAACAGTATGATCAGATACCTATAGAAATTTGTG
GGAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACCCAGATTGGCTGTACTTTAAATTTTCCAATTAGTCCTATTGAAACTGTACCAGTAA
AATTGAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAGAAAATC
AAAGCATTAACAGAAATTTGTGCAGATATGGAAAAGGAAGGAAAAATCTCAAGAATTGGACC
TGAAAATCCATACAATACTCCAGTATTTGCTATAAAGAAAAAAGACAGCACTAAATGGAGAA
AGTTAGTAGATTTCAGAGAGCTAAATAAAAGGACCCAGGACTTCTGGGAAGTTCAATTAGGA
ATACCGCATCCAGCAGGTCTAAAAAAGAAAAAATCAGTAACAATACTAGATGTGGGGGATGC
ATATTTTTCAGTTCCTTTAGATGAAAGCTTTAGAAAGTATACTGCATTCACCATACCTAGTACA
AACAATGAGGTACCAGGAATTAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGCAGCATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAATCCAG
AAATAATTATCTGTCAATACGTGGATGACTTGTATGTAGGCTCTGACTTAGAAATAGGACAAC
ATAGAGCAAAAATAGAGGAATTAAGAGAACATCTATTGAGGTGGGGGTTTACTACACCAGAC
AAAAAGTATCAGAAAGAACCTCCATTTCTTTGGATGGGATATGAACTCCATCCTGACAAATGG
Sample 17 >Busia17_A1
TGGCAACGACCCCTCGTCACAATAAAGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATGAGTTTGCCAGGAAGATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATCTGTG
GACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATC
TGTTGACTCAGATTGGTTGCACTTTAAATTTTCCCATTAGCCCTATTGAGACTGTACCAGTAAA
ATTAAAGCCAGGAATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAATAA
AAGCATTAAAAGAAATTTGTACAGAAATGGAAAAGGAAGGGCAAATTTCAAAAATTGGGCCT
GAAAATCCATACAACACTCCAATATTTGCAATAAAGAGAAAAGATAAAACTAAATGGAGAAA
AGTAGTAGATTTCAGAGAGCTCAATAAGAGAACACAAGACTTTTGGGAAGTTCAATTAGGAA

TACCGCATCCAGCGGGCTTACCAAAGAGCAAATCAGTAACAGTACTAGATGTGGGGGACGCA
TTCTTTTCAGTTCCTTTACATGAAGATTTTAGAAAATATACTGCATTCACCATACCTAGTATAA
ACAATGAGACACCAGGAATCAGATATCAGTACAATGTACTCCCACAGGGATGGAAAGGATCA
CCAGCAATATTCCAGGCTAGCATGACAAAAATCTTAGAGCCCTTTAGAGCAAAAAACCCAGA
AATGATTATCTATCAATACGTGGATGACTTGTATGTAGCRTCTGATTTGGAGATAGAACAGCA
TAGAACAAAAATAGATGAGTTAAGAAATCATCTATTGAGATGGGGATTCACTACACCAGACA
AAAAGCATCAGAAAGAACCTCCATTTCTCTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 18 >Busia18_A1
TGGCGACGACCCATCGTCACAGTAAAAGTAGGAGGACAGCTAAGAGAAGCTCTGTTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAGGTAAAACAGTATGATCAGGTAACTATAGAAATTTGTG
GAAAAAAGGTTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACTCAGATTGGCTGTACTTTAAATTTTCCCATTAGTCCTATTGAGACTGTACCAGTCA
AATTAAAGCCAGGAATGGATGGCCCAAAAATTAAACAATGGCCATTGACAGAAGAGAAAATA
AAAGCATTAACAGAAATTTGTAACGAGTTGGAAAAGGACGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAGAATAGCACTAAGTGGAGAA
AATTAGTAGATTTTAGAGAGCTCAATAAAAGGACTCAGGATTTCTGGGAAGTTCAATTAGGAA
TACCGCATCCCTCGGGCTTAAAGAAGAATAAATCAGTAACAGTACTAGATATAGGGGATGCA
TATTTTTCAGTTCCATTAGATGAAAACTTTAGAAAGTATACTGCATTCACCATACCTAGCATAA
ACAATGAGACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCA
CCAGCAATATTCCAGAGTAGCATGCTAAAGATCTTAGAGCCCTTTAGAACAAAAAATCCAGA
AATAACCATCTATCAGTACGTGGATGACTTGTATGTAGGGTCTGATCTGGAAATAGAGCAACA
TAGAGCAAAAGTAGATGAGTTAAGAGATTATCTATGGAAGTGGGGATTTTATACACCAGATA
AAAAGCATCAGAAAGAACCCCCATTTCAGTGGATGGGATATGAGCTCCATCCTGACAAATGG
Sample 19 >A1_A1
TGGCAACGACCCCTAGTCACAGTAAAGATAGGGGGACAGCTAAAGAAGCTTTATTAGATACA
GGAGCAGATGATACAGTATTGGAAGATATAAATTTGCCAGGAAAATGGAAACCAAAAATGAT
AGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAGATACCTATAGAAATTTGTGG
AAAAAGAGCTATAGGTACAGTATTAATAGGACCTACACCTGTCAACATAATTGGAAGAAATA
TGTTGACCCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCCTATTGATACTGTACCAGTAGC
ATTGAAACCAGGAATGGATGGCCCAAGGGTTAAACAATGGCCATTGACAGAAGATAAAATAA
AAGCATTAACAGAAATTTGTAAAGAAATGGAAAAGGAAGGCAAAATCTCAAGAATTGGGCCT
GAAAATCCATACAATACCCCAGTATTTGCTATAAAGAAAAAGGACAGCACTAAATGGAGGAA
ATTAGTAGATTTCAGAGAACTCAACAAAAGAACTCAGGACTTCTGGGAAGTTCAATTAGGAAT
ACCACATCCAGCAGGTTTAAAGAAGAAAAAGTCAGTAACAGTACTAGATGTGGGGGACGCAT
ACTTTTCAGTTCCCTTAGATGAAAACTTTAGAAAGTACACTGCATTCACCATACCTAGTACAA
ACAATGAGACACCAGGAATCAGATACCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCA
CCGGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATTACAAAATCCAGAA
ATAATCATCTATCAATACGTGGATGACCTGTATGTAGGGTCTGATTTAGAAATAGGGCAGCAT
AGAAAAAAGATAGAGGAGTTGAGAGCTCATCTATTGAGCTGGGGGTTTACTACACCAGACAA
AAAGCACCAGAAAGAACCTCCATTTCTTTGGATGGGATATGCATTGCATCCTGACAAATGG
Sample 20 >seq 2>B045_A1
TGGCAACGACCCCTCGTCACAATAAAGGTAGGGGGACAGCAAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAGATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAAATACCTATAGAAATTTGTG
GAAGAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGCTTGGGTGTACTTTAAATTTCCCAATTAGTCCTATTGAAACTGTACCAGTAA
AATTAAAACCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA

AAAGCATTAACAGAAATTTGTACAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCAATAAGGAAAAAAGATAGTACTAAATGGAGGA
AGCTAGTAGATTTTAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAGGTTCAATTGGGA
ATACCGCATCCAGCGGGCCTAAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGACGC
ATATTTTTCAGTTCCTTTAGATGAAAGCTTTAGAAAGTATACTGCGTTCACCATACCTAGTACA
AACAATGCAACACCAGGAATCAGATATCAGTACAATGTACTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTTCAGAGTAGCATGACAAAAATCTTAGATCCCTTTAGATCAAAAAATCCAGA
AATAATTATCTATCAATATATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGCA
TAGAGCAAAAGTAGAAGAGTTGAGAGCTCATCTATTGAGATGGGGATTTACTACACCAGATA
AAAAGCATCAGAAAGAACCCCCATTCCTTTGGATGGGTTATGAGCTCCATCCTGACAAGTGG
Sample 21>4r_D
TGGCAGCGACCCCTAGTCACAGTAAAGATAGGAGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAAATTTACCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGCTTTATCAAAGTAAGACAGTATGATCAAATTCTAGTAGAAATCTGTG
GACATAAAGCTATCGGTACAGTATTAGTAGGACCTACACCTGTCAACATCATTGGAAGAAATT
TGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTACCAGTAAA
ATTAAAGCCAGGGATGGATGGTCCAAAAGTTAAACAATGGCCATTGACAGCAGAAAAAATAA
AGGCACTAACAGAAATTTGTACAGAAATGGAGAAGGAAGGAAAAATTTCAAAAATTGGGCCT
GAAAATCCATATAATACTCCAATATTTGCAATAAAGAAAAAAGACAGTACTGAGTGGAGAAA
GTTAGTAGATTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAATTCAACTAGGAAT
ACCACATCCTGCAGGGTTAGAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCTT
ATTTTTCAGTTCCTTTATATGAAGACTTTAGAAAATATACTGCATTCACTATACCTAGTATAAA
TAATGAGACACCAGGAATTAGATATCAGTACAATGTACTTCCACAAGGATGGAAAGGATCAC
CAGCAATTTTCCAAAGTAGCATGACAAGAATCTTAGAACCTTTTAGAAAACAAAATCCAGAA
ATGATTATTTGTCAATACATAGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAGCAT
AGAATGAAAATAGAGGAATTAAGGGAACATCTGTTGAAGTGGGGATTTACCACACCAGACAA
AAAGCATCAGAAAGAACCTCCATTTCGTTGGATGGGTTATGAACTCCATCCTGATAAATGG
Sample 22>010_A1
TGGCAACGACCTCTTGTCACAGTAAAAGTAGGGGGACAGCAAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGACACAGTATTAGAAGACATAGATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAAATACCTATAGAAATTTGTG
GAAGAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGCTTGGTTGTACTCTAAATTTCCCAATTAGTCCTATTGAAACTGTACCAGTAA
AATTAAAACCAGGAATGGATGGCCCAAGGGTTAAACAATGGCCATTGACAAAAGAAAAAATA
AAAGCATTAACAGAAATTTGTACAGAGATGGAAAAGGAAGGAAAAATTTCAAGAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCAATAAAGAAAAAAGATAGCAACAAATCGGAGA
AAGCTAGTAGAACTTTAGAGAGCTCAATAAAAGAAACACAAGAACTTTTGGGAGGTTCAATT
CGGGAATACCGCATCCAGCGGGCCTACCAAAGAAAAAATCAGTAACAGTACTAGATGTGGGA
GACGCATATTTTTCAGTCTCTTTAGATGAAAGCTTTAGAAAGTATACTGCGTTTACCATACCTA
GTACAAACAATGAAACACCAGGAATCAGATATCAGTACAATGTACTTTCCACAGGGATGGAA
AGGATCACCAGCAATATTTCAAAGTAGCATGACAAAAATCTTAGATCCCTTTAGATTAAAAAA
TCCAGAAATAATTATCTATCAAATGGTGGGGAATTTGTATGTAGGATCTGATTTAGAAATAGG
GCAGCATAGAGCGAAAGTAGAAGAGTTAAGAGCTCATCTATTGAAATGGGGATTTACTACAC
CAGATAAAAAGCATCAGAAAGAACC
Sample 23 >B11_A1
TGGCAACGACCCCTTGTCACAATAAAAATAGGGGGACAAATAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTYATCAAAGTAAAACAGTATGATCAGATAACTATAGAAATTTGT
GGAAAAAAGGCTATAGGTACAGTATTAGTAGGGCCTACACCTGTCAACATAATTGGAAGAAA
CATGTTGACCCAGATTGGTTGCACTCTAAATTTCCCAATTAGTCAGATTGAGACTGTACCAGTG

AAATTGAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAAT
AAAAGCATTAACAGAAATCTGTACAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGC
CTGAAAATCCATACAATACTCCAATATTTGCAATAAAGAAAAAGGATAGCACTAAATGGAGA
AAAATAGTAGATTTCAGAGAACTCAATAAAAGAACACAAGACTTCTGGGAAGTTCAACTAGG
AATACCGCATCCAGGGGGCCTAAAAAAGAACAAATCAGTGACAGTACTAGATGTGGGGGATG
CATATTTTTCAGTACCTCTAGATGAAAGCTTTAGAAAATATACTGCATTCACCATACCTAGTCT
AAACAATGAGACACCAGGAGTCAGGTATCAGTACAATGTGCTCCCACAGGGATGGAAAGGAT
CACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAGCAAAAAATCCA
GAAATAGTTATCTACCAATATGTGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAG
CATAGAGCAAAAATAGAGGAGTTGAGAGCTCATCTATTGAGCTGGGGATTTACTACCCCAGA
CAAAAAGCATCAGAAAGAACATCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAATG
G

Sample 24seq >035_A1B

TGGCAGCGACCCCTTGTCTCAATAAAAGTAGGAGGGCAGATAAAAGAGGCTCTCTTAGACAC
AGGAGCAGATGATACTGTATTAGAAGAAATAAAATTGCCAGGAAATTGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATGAAATACCTATAGAAATCTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTGGGACCAACACCTGTCAACATAATTGGAAGAAAT
CTGTTGACTCAGCTTGGATGCACACTAAATTTTCCAATTAGTCCCATTGAAACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTCAAACAATGGCCATTGACAGAAGAAAAGATA
AAAGCATTAACAGAAATTTGTGCAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATAGGGCC
TGAAAATCCATACAATACTCCAATATTTGCAATAAAGAAAAAGGACAGTGGTAGATGGAGAA
AATTAGTAGATTTCAGAGAGCTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGAA
TACCACATCCAGCAGGTTTAGAAAAGAAAAAATCAGTAACAGTACTAGATGTAGGGGATGCA
TATTTTTCAGTTCCCTTAGATGAAAAGTTTAGAAAATATACTGCATTTACCATACTTAGTACAA
ACAATGCAACACCAGGAAGTCGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCA
CCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAGCCTTTTAGAAAACAAAATCCAGA
CATAGTTATCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAGCA
TAGAACAAAAATAGAGGAGCTGAGACAACATCTGTTGAGGTGGGGACTTACCACACCAGACA
AAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGAACTCCAACCTGATAAATGG

Sample 25seq >A3r_D

TGGCAACGACCCATTGTTACAATAAAGATAGGGGGACAATTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTGTTAGAAGATATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGCGGGAATTGGAGGTTTTGTCAAAGTAAGAGAGTATGATCAAGTACCTATAGAAATCTGTG
GACATAAAGTTATAGGTACAGTATTAGTAGGACCTACACCTTCCAACATAATTGGGAGGAATT
TGTTGACTCAGATTGGCTGCACTTTAAATTTTCCAATTAGTCCTATTGAAACTGTACCAGTAAA
GTTAAAACCAGGGATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAGAAAATAA
AAGCACTAACAGAAATTTGTACAGAGCTGGAAAAGGAAGGAAAAATTTCCAGAATTGGGCCT
GAAAATCCATACAATACTCCAATATTCGCCATAAAGAAAAAGGACAGTACTAAGTGGAGAAA
ATTAGTAGATTTCAGAGAACTTAATAAGAGAACTCAAGATTTCTGGGAAGTTCAACTAGGAAT
ACCACATCCTGCAGGGCTAAAAAAGAAAAAATCAGTAACAGTATTGGACGTGGGTGATGCAT
ATTTCTCAGTTCCCTTAGATGAAGACTTTAGAAAGTATACTGCATTCACCATACCTAGTACAAA
CAATGAGACACCAGGGATTCGATATCAGTACAATGTGCTTCCACAAGGATGGAAAGGATCAC
CAGCAATATTCCAAAGCAGCATGACAAAAATCTTAGAACCTTTTAGAAAACAAAATCCAGGA
ATAATTATCTATCAATACGTGGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGAAACAT
AGAGAAAAAATAGAGGAACTAAGGGAACATCTATTGAAGTGGGGATTTTACACACCAGACAA
AAAACATCAGAAAGAACCTCCATTTCTTTGGATGGGTTATGAACTCCATCCTGATAAATGG

Sample 26>A65_A1

TGGCAACGACCCCTTGTTACAGTAAAAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGAGGAATTGGAGGTTTCATCAAAGTAAAACAGTATGATCAGATACTTATAGAAATTTGTG
GGAAAAAAGCTATAGGTACAGTATTAGTAGGACCCACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTACCAGTAA

AACTAAAACCAGGAATGGATGGCCCAAAGGTGAAACAATGGCCATTGACAGAAGAAAAAAT
ACAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGGAAAATTTCAAGAATTGGGC
CCGAAAATCCATACAATACTCCAATATTTGCTATAAAGAGAAAAGAAAGCACTAAATGGAGG
AAATTAGTAGATTTCAGAGAGCTCAATAAAAGAACTCAAGATTTTTGGGAAGTTCAATTAGGA
ATACCGCATCCAGCGGGCTTATTGAAGAAGAAATCAGTAACAGTACTAGATGTGGGGGACGC
ATATTTTTCAGTTCCTTTAGATGAAAGTTTTAGGAAATATACGGCRTTCACCATACCTAGTACA
AACAATGAGACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATTTTAGAGCCCTTCAGATCAAAAAATCCAG
AAATAATTATCTATCAATACGTGGATGACCTGTTGGTAGGGTCTGATTTAGAAATAGGGCAAC
ATAGAACAAAAATAGAGGAGTTAAGAGCTCATCTGTTRAAATGGGGATTTACTACCCCAGAC
AAAAAGCATCAGAAAGAACCACCTTTTCTTTGGATGGGGTATGAACTCCACCCTGACAAGTGG

Sample 27 >A6r_D

TGGCAACGACCCCTCGTCACAGTAAAGATAGGGGGACAGCTAAAGGAAGCTCTATTAGACAC
AGGAGCAGATGATACAGTATTAGAAGAAATGAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGCTTTATCAAAGTAAGACAGTATGAGCAAATACCAGTAGAAATCTGT
GGACATAAAGCTATAGGTACAGTATTAATAGGACCTACACCTGTCAACATAATTGGGAGAAA
TTTGTTGACTCAGATTGGCTGTACTTTAAATTTTCCAATTAGTCCTATTGAAACTGTACCAGTA
AAATTGAAGCCAGGGATGGATGGTCCAAAAGTTAGACAATGGCCATTGACAGAAGAGAAAAT
AAAAGCATTAACAGAAATTTGTCTGGAAATGGAAAAGGAAGGAAAGATTTCAAAAATAGGAC
CTGAAAACCCATACAATACTCCAATATTTGCCATAAAGAGAAAGGACAGTACTAAATGGAGA
AAATTAGTAGATTTCAGAGAACTTAATAAGAGAACTCAAGACTTTTGGGAAGTTCAATTAGGA
ATACCACATCCTGCAGGGCTAAAAAAGAATAAGTCAGTAACAGTACTGGATGTGGGTGATGC
ATTTTTTTCAGTTCCCTTATATGAAGACTTTAGAAAATATACAGCATTCACCATACCTAGTATA
AATAATGAAACACCAGGAATTAGATATCAGTACAATGTGCTCCCACAAGGATGGAAAGGATC
ACCGGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAACCTTTTAGAAAACAAAACCCAG
AAATGTTTATCTGTCAATACATGGATGATTTGTATGTAGGGTCTGACTTAGAAATAGGGCAGC
ATAGAGTAAAAATAGAGGAATTAAGGGAACACCTATTAAAGTGGGGATTTACCACACCAGAC
AAAAAGTATCAAAAAGAACCTCCATTTCTTTGGATGGGTTATGAGCTCCATCCTGATAAATGG

Sample 28>047_A1B

TGGCAACGACCCCTCGTCACAATAAAGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATGAGTTTGCCAGGAAGATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTCATCAAAGTAAAACAATATGATCAGATAGTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTGGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATATTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGACACTGTACCTGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTAACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTACAGAGATGGAAAAGGAAGGAAAAATCTCAAAAATAGGGC
CTGAAAATCCATATAATACTCCAATATTTGCTATAAAGAAAAAAGACAGCACTAAATGGAGA
AAATTGGTAGATTTCAGAGAGCTCAATAAAAGAACTCAAGACTTTTGGAAAGTTCAATTAGGG
ATACCACATCCAGCAGGCTTAAAAAAGAACAAATCAGTAACAGTACTAGATGTGGGGGATGC
ATATTTTTCAGTTCCCTTAGATGAAAACTTTAGAAAGTATACTGCATTCACCATACCTAGTACA
AACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAACCCTTTAGATCAAAAAATCCAG
AATTAATTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGACAGC
ATAGAACAAAAATAGAAGAGCTAAGAGCTCATCTATTGAGCTGGGGACTTACTACACCAGAC
AAAAAGCATCAGAAAGAACCTCCCTTTCTTTGGATGGGATATGAACTCCATCCTGACAAGTGG

Sample 29>049_A1J

TGGCAGCGACCCCCTGTCACAGTAACAGTAGGCGGGCACCTAAAAGAAGCTCTTTTACATACA
CGAGCAGATGATCCAGTATTAGAAGAGATAGATTTGCCAGGAAAATGGAAACCAAAAATGAT
AGGGGGAATYGGAGGTTTTATCAAGGTAAAACAGTATGATCAGATACTTATAGAAATTTGTG
GAAAAAGGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACTCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTGCCAGTAA
AATTAAAGCCAGGAATGGATGGCCCTAGGGTTAAACAATGGCCATTGACAGAAGAAAAAATA

AAAGCATTAACAGAAATTTGTATGGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCAATAAGGAAAAAGGATAGCACTAAATGGAGAA
AATTAATAGATTTCAGAGAACTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAACTAGGA
ATACCGCATCCAGCGGGTTTAAAAAAGAAAAAATCAGTAACAGTTCTAGATGTGGGAGACGC
ATATTTTTCAGTTCCCTTAGATGAAAGTTTTAGAAAGTATACAGCGTTCACCCTACCTAGTACA
AACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCATTTAGATCAAAAAATCCAG
AGATAATTATCTGCCAATACGTGGATGATTTGTATGTAGGATCTGATTTAGAAATAGAGCAGC
ATAGAGCAAAAATAGAAGAGTTAAGGGCTCATTTATTGAGCTGGGGACTTACTACCCCAGAC
AAAAAGTATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAACTCCATCCTGACAAATGG

Sample 30>A72_D

TGGCAACGACCCCTCGTCACAATAAAGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATGAATTTGTCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAATATGAACAAATACCTGTAGAAATTTGTG
GACATAAAGCTATAGGTACAGTGTTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATT
TGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTCCTATTGACACAGTACCAGTAAA
ATTAAAGCCAGGGATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAATAA
AAGCACTAATAGAAATTTGTACAGATATGGAAAAGGAAGGAAAAATTTCAAGAATTGGGCCT
GAAAATCCATATAATACTCCAATATTTGCCATAAAGAAAAAAGACAGTACTAGGTGGAGAAA
ATTAGTAGATTTCAGGGAACTTAATAAAAGAACTCAAGACTTCTGGGAAGTTCAACTAGGAAT
ACCACATCCCGCAGGGCTAAAAAAGAACAAATCAGTAACAGTATTGGATGTGGGTGACGCAT
ATTTCTCAGTTCCTTTATATGAGGACTTTAGAAAATATACTGCATTCACCATACCTAGTTTGAA
CAATGAGACACCAGGAATTAGATATCAGTACAATGTGCTTCCACAAGGATGGAAAGGATCAC
CAGCAATATTTCAAAGTAGCATGACAAAAATCTTAGACCCTTTTAGAAAACAAAATCCAGACA
TAGTTATTTATCAATACGTGGATGATTTGTATGTAGGATCTGACTTGGAAATAGGGCAACATA
GAGCAAAAATAGAGGAATTAAGGGAACACCTCTTGAAATGGGGGTTTACCACACCAGACCAA
AAGCATCAAAAAGAACATCCATTTCATTGGATGGGCTATGAACTCCATCCTGATCAATGG

Sample 31>043B_A1

TGGCAACGACCTCTTGTCACAGTAAAAATAGGGGGACAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTCATCAAGGTAAAACAGTATGATCAGATACTTATAGAAATTTGTG
GGAAAAAGGCTATAGGTACAGTATTGGTAGGACCTACACCTGTCAACATAATTGGGAGAAAT
ATGTTGACCCAGATTGGTTGTACTCTAAATTTCCCAATTAGTCCTGTTGCTACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGTCCAAGGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCTATAAAGAGAAAAGACAGCACTAAATGGAGAA
AATTAGTAGATTTCAGAGAGCTCAATAAAAGAACTCAAGACTTCTGGGAAGTTCAATTAGGA
ATACCGCATCCAGCGGGCTTAGAAAAGAAAAAATCAGTGACAGTACTAGATGTGGGGGATGC
ATATTTTTCAGTTCCTTTAGATGAAAGCTTTAGAAAGTATACAGCATTCACCATACCTAGTACA
AACAATGAGACACCAGGAATTAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCGGCAATATTCCAGAGTACCATGACAAAAATCTTAGAGCCCTTTAGATCACAAAATCCAGA
CATAATTATCTGTCAATACGTGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGCA
TAGAGCAAAGGTAGAGGAGTTAAGAGATCATCTATTGAGCTGGGGATTGACTACACCAGACA
AAAAGCATCAGAAAGAACCTCCATTTCTTTGGATGGGATATGAACTCCATCCTGACTAGTGG

Sample 32>066_A1

TGGCAACGACCCATCGTCACAGTAAGGATAGGGGGACAGCTAATAGAAGCTCTCTTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAGATTTGCCAGGAAAATGGACACCAAAAATGA
TAGTGGGAATTGGAGGATTAGTAAAAGTAAAACAGTATGATCAAGTACCTATAGAAATTTGT
GGAAAAAGGGCTGTAGGTGCAGTATTAGTAGGGCCTACACCTGCCAACATCATTGGAAGAAA
CATGTTGACTCAGATTGGTTGTACTCTAAATTTCCCAATTAGCCCTATTGAAACTGTACCAGTA
AAATTAAAGCCAGGAATGGATGGCCCAAGGGTTAAGCAATGGCCATTGACAGAGGAAAAAAT
AAAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAGATTGGGC

CAGAAAATCCATATAATACTCCAATATTTGCTATAAAGAAGAAAAATAGCAATAAATGGAGA
AAATTAGTAGATTTTAGAGAGCTCAATAAAAGAACACAGGACTTTTGGGAAGTTCAATTGGGC
ATACCGCATCCAGCGGGTCTAAARAAGAAAAAATCAGTAACAGTCCTAGATGTGGGTGATGC
ATATTTTTCAGTACCTTTAGATGAAAACTTTAGAAAATACACTGCATTCACCATACCTAGTATA
AACAATGAGACACCAGGAATCAGATATCAATACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCTTTTAGAGCAAAAAATCCAG
AAATAATTATCTATCAATACGTGGATGACTTGTATGTAGCATCTGATTTAGAAATAGGGCAGC
ATAGAGCAAAAATAGAAGAGTTAAGAGCTCATTTATTGAACTGGGGATTTTTTACCCCAGATG
AGAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 33>48_D
TGGCAACGACCCCTCGTCACAGTAAAGATAGGGGGACAGCTAAAGGAAGCTCTCTTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATGAATCTGCCAGGAAAATGGAAACCAAAAATG
ATAGGGGGAATTGGAGGCTTTATCAAAGTAAGACAATATGAACAAATACCTGTAGAAATTAG
TGGACATAAAGCTATAGGTACAGTGTTAGTAGGACCTACACCTGTCAACATAATTGGAAGAA
ATTTGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTCCTATTGACACAGTACCAGT
AAAATTAAAGCCAGGGATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAA
ATAAAAGCACTAATAGAAATTTGTACAGATATGGAAAAGGAAGGAAAAATTTCAAGAATTGG
GCCTGAAAATCCATATAATACTCCAATATTTGCCATAAAGAAAAAAGACAGTACTAGGTGGA
GAAAATTAGTAGATTTCAGGGAACTTAATAAAAGAACTCAAGACTTCTGGGAAATTCAACTA
GGAATGCCACATCCTGCAGGGCTAAAAAAGAACAAATCAGTAACAGTATTGGATGTGGGTGA
TGCATATTTCTCAGTTCCTTTATATGAGGACTTTAGAAAATATACTGCATTCACCATACCTAGT
CTGAACAATGAGACACCAGGAATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGG
ATCACCAGCAATATTTCAAAGTAGCATGACAAAAATCTTAGACCCTTTTAGAAAACAAAATCC
AGACATAGTTATTTATCAATACGTGGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCA
GCATAGAGCAAAAATAGAGGAATTAAGGGAACACCTCTTGAAATGGGGGTTTACCACACCAG
ACCAAAAGCATCAGAAAGAACATCCATTTCACTGGATGGGCTATGAACTCCATCCTGATCAAT
GG
Sample 34>071_D
TGGCAACGACCCCTTGTCACAGTAAAAATAGGGGGACAACTAAAAGAAGCTCTGTTAGATAC
AGGAGCAGATGACACAGTATTAGAAGAAATAAATTTGCCAGGAAAATGGAAGCCAAAAATG
ATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAAATTCCTGTAGAAATCTGT
GGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAA
TTTGTTGACTCAGATTGGGTGCACTTTAAATTTTCCAATTAGTCCTATTGAAACTGTACCAGTA
AAATTAAAGCCAGGGATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAAT
AAAAGCACTAATAGAAATTTGTGCAGAACTGGAAAAGGAGGGGAAAATTTCAAGAATTGGGC
CTGAAAACCCATATAATACTCCAATATTTGCCATAAAGAAAAAGGACAGCACTAAGTGGAGA
AAATTAGTAGATTTTAGAGAACTTAATAAAAGAACTCAGGATTTCTGGGAAGTTCAACTAGGA
ATACCACACCCTGCAGGGCTAAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATGC
ATATTTTTCAGTACCCTTAGATGAAGACTTTAGAAAATACACTGCATTTACCATACCTAGTATA
AACAATGAGACACCGGGAGTTAGATATCAGTACAATGTGCTCCCACAAGGGTGGAAAGGATC
CCCGGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAACCTTTTAGAAAACAAAATCCAG
ACATAGTTATCTATCAATACGTGGATGATTTGTTAGTGGGATCTGACTTAGAAATAGGACAGC
ATAGAATAAAAGTAGAGGAATTAAGGGAACACCTATTGAAGTGGGGATTTTTCACACCAGAC
AAAAAGCATCAGAAAGAACCTCCATTTCTGTGGATGGGTTATGAACTCCATCCTGATAAATGG
Sample 35>082_A1A2
TGGCAACGACCCGTTGTCACAGTAAAAGTAGGGGGACAACTAAGAGAAGCTCTGTTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAGATTTGCCAGGAAAATGGAAGCCAAAAATGA
TAGGGGGAATTGGAGGATTCATCAAAGTAAAACAGTATGATCAAATACTTATAGAAATCTGT
GGAAAAAAGGCTATAGGTACAGTATTAGTAGGGCCTACACCCGTCAACATAATTGGAAGAAA
CATGTTGACCCAGATTGGTTGTACTTTAAACTTCCCCATTAGTCCTATTGAGACTGTACCAGTA
AAATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAAT
AAAGGCATTAACAGAAATTTGTACAGAGATGGAGAAGGAAGGAAAAATCTCAAAAATTGGGC

CTGAAAATCCATACAATACTCCCATATTTGCAATAAAGAGAAAAGATGGCACTAAATGGAGA
AAATTAGTAGATTTCAGAGAGCTCAATAAAAGAACACAGGACTTTTGGGAAGTTCAATTAGG
AATACCACATCCAGCAGGTCTAAAAAAGAATAAATCAGTAACAGTACTGGATGTGGGGGACG
CATATTTTTCAGTTCCTTTAGATGAAGACTTTAGGAAGTATACTGCATTCACCATACCTAGTAT
AAACAATGAGACACCAGGGATTAGATATCAGTACAATGTGCTTCCACAGGGATGGTTAGAAA
CAGCAGCAATGTTCCAAAGTTGCAAGACAAAGATTTTACAGGAAAATAGAAGACAAAAACGA
TAGGGGGAAATCGGAGGATTCATGAAAGTAAAACAGTATGATCAAATATTTATAKAGGGCAG
CATAGAACAAAAATAGAGGAGCTGAGACAACATCTGTTGAGGTGGGGACTTACCACACCAGA
CAAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGAACTCCATCCTGATAAATG
G
Sample 36>145_A1
TGGCAACGACCCATTGTCACAGTAAAAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTRTTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGAGGAATTGGGGGTTTCATCAAGGTAAAACAGTATGATCAAATAATGATAGAAATTTGT
GGAAAAAAGGCTATAGGCACAGTGTTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAA
CATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTGCCAGTA
AAATTAAAGCCAGGAATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAGGAAAAAAT
AAAAGCATTAACAGAAATTTGTACAGATATGGAAAAGGAAGGAAAAATTTCAAGAATTGGGC
CTGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGACAGCACTAAATGGAGG
AAATTAGTAGATTTCAGAGAGCTCAATAAAAGAACACAAGATTTCTGGGAAGTTCAATTAGG
AATACCACATCCAGCGGGCTTAAAAAAGAACAAATCAGTAACAGTACTAGATGTGGGGGACG
CATATTTTTCAGTTCCTTTAGATGAAGAGTTTAGAAAATATACTGCATTCACCATACCTAGTAC
AAACAATGAGACACCAGGAACCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGAT
CACCGGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATTAAAAAATCCAG
AGATAATTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGC
ATAGAACAAAAATAGAAGAGTTGAGAGCTCATCTATTGAGCTGGGGATTTACTACACCAGAC
AAAAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 37>90r_D
TGGCAACGACCCCTCGTCACAGTAAAGATAGGGGGACAGCTAAAGGAAGCTCTCTTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATGAATTTGTCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGCTTTATCAAAGTAAGACAATATGAACAAATACCTGTAGAAATTTGT
GGACATAAAGCTGTAGGTACAGTGTTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAA
TTTGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTCCTATTGACACAGTACCAGTA
AAATTAAAGCCAGGGATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAAT
AAAAGCACTAATAGAAATTTGTACAGATATGGAAAAGGAAGGAAAAATTTCAAGAATTGGGC
CTGAAAATCCATATAATACTCCAATATTTGCCATAAAGAAAAAAGACAGTACTAGGTGGAGA
AAATTAGTAGATTTCAGGGAACTTAATAAAAGAACTCAAGACTTCTGGGAAGTTCAACTAGG
AATACCACATCCTGCAGGGCTAAAAAAGAACAAATCAGTAACAGTATTGGATGTGGGTGATG
CATATTTTTCAGTTCCTTTATATGAGGATTTTAGAAAATATACTGCATTCACCATACCTAGTCT
GAACAATGAGACACCAGGAATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGAT
CACCAGCAATATTTCAAAGTAGCATGACAAAAATCTTAGACCCTTTTAGAAAACAAAATCCAG
ACATAGTTATTTATCAATACGTGGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAGC
ATAGAGCAAAAATAGAGGAATTRAGGGAACACCTCTTGAAATGGGGATTTACCACACCAGAC
CAAAAGCATTCAGAAAGAACATCCATTTCATTGGATGGGTTATGAACTCCATCCTGATCAATG
G
Sample 38 >Busia38_D
TGGCAACGACCCCTTGTCACAGTAAAGATAGGGGGACAACTAAAGGAAGCTCTCTTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATGAATTTGCCAGGGAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGCTTTATCAAAGTAAGACAGTATGAGCAAATACTTGTAGAAATCTGT
GGACATAAAGCAATAGGTACAGTATTAATAGGACCTACACCTGTCAACATAATTGGAAGAAA
TTTGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTACTATTGAAACTGTACCAGTA
AAATTAAAGCCAGGGATGGATGGCCCAAAAGTTAAACAGTGGCCATTGACAGAAGAAAAAAT

AAAAGCACTGATAGAAATCTGTACAGAAATGGAAAAGGAAGGGAAAATTTCAAGAATTGGGC
CTGAAAATCCGTACAATACTCCAATATTTGCCATAAAGAAAAAAGACAGTATTAAGTGGAGA
AAATTAGTAGATTTCAGGGAACTTAATAAGAAAACTCAAGACTTCTGGGAAATTCAGCTAGG
AATACCACATCCTGCAGGACTAAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATG
CATATTTTTCAGTTCCCTTATATGAAGAATTTAGAAAATATACTGCATTCACCATACCTAGTGT
AAACAATGAGACACCAGGAATTAGATATCAGTACAATGTGCTTCCACAAGGATGGAAAGGAT
CACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAACCTTTTAGAAAACAAAACCCA
GAGATGGTTATCTATCAATACATGGATGACTTGTATGTAGGATCTGACTTAGAGATAGGGCAA
CATAGAGTAAAAATAGAGGAGTTAAGAGAACACCTCTTGAAGTGGGGGTTTACCACACCAGA
CAAAAAGCATCAGAAAGAACCTCCATTTCTTTGGATGGGTTATGAACTCCATCCTGATAAATG
G

Sample 39 >Busia39_A1C

TGGCAGCGACCCCTTGTCACAATAAAAGTAGGGGGACAGGTGAAGGAAGCTCTCTTAGATAC
AGGAGCAGATGATACTGTATTAGAAGAAATAAAATTACCAGGAAATTGGAAACCAAAAATGA
TAGGAGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAAATAGTTATAGAGATTTGTG
GAAAAAAGGCCATAGGGTCAGTGCTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACTCAGCTTGGATGCACACTAAATTTTCCAATTAGTCCTATAGAAACTGTACCAGTA
AAATTGAAACCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAAT
AAAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGGAAAATTTCAAGAATTGGGC
CTGAAAACCCATACAACACTCCAATATTTGCAATAAAGAAAAAAGATAGCACTAAATGGAGG
AAATTAGTAGATTTCAGAGAACTTAATAAAAGAACTCAAGATTTTTGGGAAGTTCAATTAGGA
ATACCGCATCCAGCAGGCTTAAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGACGC
ATATTTTTCAGTTCCCTTAGATGAAAGCTTTAGAAAATACACTGCATTCACTATACCCAGCATA
AACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATTAAAAAATCCAGA
CATAGTTATTTACCAATACATGGATGATTTGTATGTAGGATCTGATCTAGAAATAGGACAGCA
TAGAACAAAAATAGAAGAGTTGAGAGATCATCTCTTGAGATGGGGATTCACTACCCCAGACA
AAAAGCATCAGAAAGAACCTCCATTCCTGTGGATGGGGTATGAACTCCATCCTGACAAATGG

Sample 40>Busia40_A1

TGGCAACGACCCCTCGTCACAGTAAAGATAGAAGGACAGCTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAGATTTGCCAGGAAAATGGAAACCAAGAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTTATAGAAATTTGTG
GAAAAAGGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACTCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTACCAGTAA
CATTAAAGCCAGGAATGGATGGCCCAAGAATTAAACAATGGCCCTTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTCAAGAGATGGAAAAAGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGATAGCACCAAATGGAGAA
AATTAGTAGATTTTAGAGAGCTCAATAAAAGAACTCAGGATTTCTGGGAGGTGCAATTAGGA
ATACCACACCCAGCGGGTTTAAAACAGAGAAAATCAGTAACAGTACTAGATGTGGGGGATGC
ATATTTTTCAGTTCCCTTAGATGAAAACTTTAGAAAGTATACTGCATTCACCATACCTAGTATA
AACAATGAGACGCCAGGAATCAGATATCAGTACAATGTACTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATTTTAGAGCCCTTCAGAGCACAAAATCCAG
AAATAGTTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGC
ACAGAGCAAAAGTAGAGGAGTTGAGAGCTCATCTATTGAAGTGGGGATTTACCACACCAGAC
AAAAAGCATCAGAAGGAACCTCCATTTCTTTGGATGGGATATGAGCTCCATCCTGACAAATGG

Sample 41 >Busia41_A1B

TGGCAACGACCAGTTGTCACAGTAAAAATAGAGGGACAATTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTACTAGAAGATATAAATTTGCCAGGGAAGTGGAAACCAAAAATGA
TAGGGGGAATTGGGGGTTTCATCAAAGTAAGACAGTATGATCAAGTAGTTTTAGAAATTTGTG
GAAAAAAGGCTATAGGTTCAGTACTGGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACCCAGATTGGCTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTAACAGAAGAGAAAATA

AAAGCATTAACAGAAATTTGTACAGAAATGGAAAAAGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCCTACAATACTCCAATATTTGCTATAAAGAAAAAAGACAGCACTAAATGGAGAA
AATTAGTAGATTTCAGAGAGCTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGGA
TACCGCATCCAGCGGGCCTAAAAAAGAAAAAATCAGTTACAGTACTAGATGTGGGGGACGCA
TATTTTTCAGTTCCTTTAGATGAAAGCTTTAGGAAATATACTGCTTTCACCATACCTAGTTTAA
ACAATGAGACACCAGGAATTAGGTATCAGTATAATGTGCTCCCACAAGGGTGGAAAGGGTCA
CCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAACCTTTTAGAAAGAAAAATCCAGA
AATAGCTATCTATCAATACATGGATGACTTGTATGTAGGATCTGACTTAGAAATAGGGCAGCA
CAGAATAAAAATAGAGGAATTAAGAGAACACCTGTTAAAGTGGGGATTTACTACACCAGACA
AAAAACATCAGAAGGAACCTCCATTCCTTTGGATGGGATATGAACTCCATCCTGATAAATGG
Sample 42>Busia42_A1
TGGCAACGACCCCTTGTCACAGTAAAAGTAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGCTTTATCAAGGTAAAACAGTATGATCAGATAACTATGGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTACTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGCTTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AGAGCATTAACAGAAATTTGTACAGATATGGAAAAAGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCAATAAAGAAAAAAGATAGCACTAAATGGAGAA
AATTAGTAGACTTCAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGGA
ATACCCCACCCAGCAGGCCTAAAAAAGAAAAAGTCAGTAACAGTACTAGATGTAGGGGATGC
ATATTTTTCAGTTCCCTTAGATGAAGACTTCAGGAAGTATACTGCATTCACCATACCTAGTATA
AACAATGAGACACCAGGAATCAGATATCAATACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCGGCAATATTCCAGAGTAGCATGACAAAAATTTTAGAGCCTTTTAGAGCACAACATCCAGA
GATAATTATCTATCAATACATGGATGACTTATATGTAGGATCTGATTTAGAAATAGAGCAGCA
TAGAACAAAAATAGCAGAGTTAAGAGCTCATCTGTTGAGCTGGGGATTCACCACACCAGACA
AAAAACATCAGAAAGAACCCCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 43 >Busia43_D
TGGCAACGACCCCTCGTCACAGTAAAGGTAGGGGGACAATTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATGAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCATATACTAATAGAAATCTGTG
GACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTTAACATAATAGGAAGAAATT
TGCTGACTCAGATTGGTTGCACTCTAAATTTTCCAATTAGTCCTATTAAAACTGTACCAGTAAA
ATAAAAGCCAGGGATGGATGGCCCGAGAGTTAAACAATGGCCATTGACAGAAGAAAAAATA
GAAGCACTAATAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAGAATTGGGCC
TGAAAATCCATATAATACTCCAATATTTGCCATAAAGAAAAAAGACAGTACTAAGTGGAGAA
AATTAGTAGATTTCAGGGAACTTAATAAAAGAACTCAAGACTTTTGGGAAGTTCAACTAGGGA
TACCGCATCCGGCAGGGCTAAAAAAGAAAAAATCAGTAACAGTATTGGATGTGGGTGATGCA
TATTTTTCAGTCCCTTTATATGAAGACTTTAGAAAATATACAGCATTCACCATACCTAGTAGAA
ATAATGAGACACCAGGAATTAGGTATCAGTACAATGTGCTTCCACAAGGATGGAAGGGATCA
CCAGCAATATTCCAAAGTAGCATGACGAAAATCTTAGAACCTTTTAGAAAACAAAATCCAGA
AATGATTATCTATCAATACATGGATGATCTGTATGTAGGATCTGACTTAGAAATAGGGCAGCA
TAGAACAAAGATAGAGGAATTAAGGAAACATTTATTGAAGTGGGGATTTACCACACCAGACA
AAAAGCATCAGAAAGAACCTCCATTTCTTTGGATGGGGTATGAACTCCATCCTGATAAATGG
Sample 44 >A9r_A1
TGGCAACGACCCCTTGTCACAGTAAGAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTAGTAGAAGATATAAATTTGCCAGGAAAATGGAAACCAAGAATGA
TAGGGGGCATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAGATACATATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTACTAGTAGGACCTACACCTGTTAACATAATTGGAAGAAAT
ATGTTAACTCAGATTGGCTGTACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTACCAGTAA
CACTAAAGCCAGGAATGGATGGCCCAAGGATTAAACAATGGCCCTTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTAGTGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGACC

TGAAAATCCATACAATACTCCAGTATTTGCTATAAAGAAAAAAGATAGCACTAAATGGAGAA
AATTAGTAGATTTCAGGGAGCTCAATAAAAGAACTCAGGACTTCTGGGAAGTTCAATTAGGA
ATACCACATCCCGCAGGTTTAAAAAAGAAAAAATCAGTGACGGTACTGGATGTGGGGGATGC
ATATTTTCCAGTACCTTTAGATATAAACTTTAGAAAGTATACTGCATTCACCATACCTAGTATA
AATAATGAGACACCAGGAATAAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAAACTTAGAGCCCTTTAGATCAAAAAATCCAG
AAGTAATTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGC
ATAGAACAAAAGTAGAGGAGTTGAGAGCTCATCTATTGAGTTGGGGATTCACTACACCAGAC
AAAAAGCATCAGAAAGAACCTCCCTTTCTTTGGATGGGTGATG

Sample 45>03r_A1

TGGCAACGACCCCTTGTCACAGTAAAAATAGGGGGGCAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAGATTTGCCAGGGAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGGTTCATCAAGGTAAGACAGTATGATCAAATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACTCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGCGACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTACAGATATGGAAAAAGAGGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAGTATTTGCAATAAAGAAAAAGGATAGCACTAAATGGAGAA
AATTAGTAGATTTCAGAGAGCTCAATAAAAGAACACAAGATTTTTGGGAAGTTCAATTAGGG
ATACCGCATCCAGCGGGCCTAAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGACGC
ATACTTTTCAGTTCCTTTAGATGTAAGCTTTAGAAAATATACTGCATTCACCATACCTAGTATA
AATAATGAGACACCAGGAATAAGGTATCAGTACAATGTGCTTCCGCAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTCAGAGCAAATAATCCAG
AAATAATTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATCTAGAAATAGGGCAAC
ATAGAGCAAAAATAGAAGAATTAAGAGCACATCTGTTGAGCTGGGGATTTACTACACCAGAC
AAAAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGAACTCCATCCTGATAAATGG

Sample 46>02r_A1

TGGCAACGACCTGTAGTCACAGTAAGAATAGAGGGACAACTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTGGTGGGACCTACACCTGTCAACATAATTGGAAGGAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCCATTGATACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGCACAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGACAGCACTAAATGGAGAA
AATTAGTAGATTTTAGAGAGCTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGGA
TACCGCATCCAGCGGGTTTGAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGACGCC
TATTTTTCAGTTCCCTTAGATGAAGACTTTAGAAAATATACAGCATTCACCATACCTAGTACAA
ACAATGAGACACCAGGAATAAGGTATCAGTACAATGTACTTCCACAGGGATGGAAAGGATCA
CCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAGCAAAAAATCCAGA
AATGGTAATCTATCAGTACATGGATGACTTGTATGTAGGCTCTGATTTAGAAATAGGGCAGCA
TAGAGCAAAAGTAGAAGAATTAAGAAAACATCTATTGAGCTGGGGATTTACTACACCAGACA
AAAAGCATCAGAAAGAACCTCCATTTCTCTGGATGGGGTATGAACTCCATCCTGATAAGTGG

Sample 47>Busia47_A1

TGGCAACGACCCCTTGTCACAGTAAGAATAGGGGGACAGTTCAGGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAGATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGCTTCATCAAGGTAAAACAGTATGATCAGATACTTATAGAAATCTGTG
GAAAAAAGGCTATAGGTACAGTATTGGTAGGACCTACACCTGTCAACATCATTGGAAGAAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCCTATTAAGACTGTACCAGTGC
AATAAAGCCCAGGAATGGATGGCCCAAAAGTTAAACAATGGCCATTGACGGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTACAGATATGGAAAAGGAAGGGAAAATTTCAAGAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCGATAAAGAAAAAGGATAGCACTAAATGGAGGA

AATTAGTAGACTTCAGAGAGCTCAATAAAAGAACACAAGATTTTTGGGAAGTTCAATTAGGC
ATACCACATCCAGCGGGCCTCAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGATGC
ATATTTTTCAGTTCCTTTGCATGAAAGCTTTAGAAAATATACTGCATTCACCATACCTAGTATA
AACAATGAGACACCAGGAATCAGATATCAGTACAATGTGCTTCCCCAGGGATGGAAAGGATC
ACCGGCAATATTCCAGAGTAGCATGACAAACATTTTAGAGCCCTTTAGATCAAAAAATCCAGA
AATAGTTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGCA
TAGAACCAAAATAGAAGAGTTGAGAGCTCATCTATTGAGCTGGGGATTTACTACACCAGACA
AAAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGATAAGTGG
Sample 48>034_A1A2
TGGCAACGACCCCTTGTCACAGTAAAAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAAATTTGCCAGGAATATGGAAACCAAGAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAGATACCTATAGAAATTTGTG
GAAAAAAGGCTATCGGTACAGTACTAGTAGGGCCTACACCTGTTAACATAATTGGAAGAAAT
ATGTTGACTCAGATTGGTTGTACTTTAAACTTTCCAATTAGTCATATTGAGACTGGACCATAAC
CGTAACATTAAAGCCCGGATGGGACGGCCCCTAAACAATGGCCAGTGACTTAAGAAAAAAAA
AAATAAAAGCCTGAAAAGAAATTCGTGCAGAAATCGAAAAAGGAAAAAAAAATTCAAAAAT
CGGGAAATCCAAACCCATAAAATAACCCAATAATCAATAAAAAAAAAAAAGAACGGCACTA
AATCGAGAAAATCAGTAAGGTGCTCAAGAAAAGAACACAAGACTACCGGGAAGTTCGGAATT
TCATTCAACATACAGCGGGCCTAAAAAAGAAAAAATCAGTAACAGTACTGAGCGTGGGGGGC
GCATGATTTTCAGTTCCTTTATATGAAAGCTGTAGAAAATAAACTGCATACACCATACCTAGT
CTAAACAATGAAACACCAGGAAATAAGAATCAGTACAATGTGCTCCCACAGGGAAGGAAAGG
GTCACCGGCAATAACCCAGTGTAGCATGACAAAAACAATTCCGAGCTTCAGGTCAAAAAATA
CNGAACCCGTTAATTCTCAATACGTGGACGNTTCGTACGTAGGGGCAGCGTTAAAAATATAGC
AGCACAGAGCAAGGCAACAGGANAAGTAACTTTACAGGTCCAGTTGGGGAATTTTAGCACAC
CAGACCAAACATCAGAAAGAGAATTTTCACCCCTTCGGACGGGACGCGATCMCTCTCCCGAA
CAAG
Sample 49>A4r_A1
TGGCAACGACCTATTGTCACAGTAAGAATAGAGGGACAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGACACAGTATTAGAAGACATAAATTTGCCAGGGAAATGGAAACCAAGAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAGATAAGTATAGAAATTTATG
GAAAAAAGGCTATGGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTGCCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAGAAAATA
AAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATCACAAAAATTGGGC
CTGAAAATCCATACAATACTCCAATATTTGCAATAAAGAAAAAAGACAGCACTAAATGGAGA
AAATTAGTAGATTTTAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGG
AATACCACATCCAGCGGGCCTAAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGATG
CATATTTTTCAGTCCCTTTACATGAAGAATTTAGAAAGTATACTGCATTCACCATACCTAGTAC
AAACAATGAGACACCAGGAATCAGGTATCAGTATAATGTGCTTCCACAGGGATGGAAAGGCT
CACCAGCAATATTTCAGAGTAGCATGACAAAAATCTTAGAGCCTTTTAGAATAAAAAATCCAG
AAATAATTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAAC
ATAGAGCAAAAGTAGAGGAGTTAAGAGCCCATCTATTGAGCTGGGGATTTACTACACCAGAT
AAAAAGCATCAAAAAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 50>A11r_A1
TGGCAACGACCTCTTGTCACAGTAAAAATAGGGGACCAACTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAGATTTGCCAGGGAGATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAAATACCTATAGAAATTTGTG
GAAAAAGGGCTATAGGTACAGTGTTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTACCAGTAA
AATTGAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTACAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCCATAAAGAAAAAGGATAGCTCTAAATGGAGAA

AATTAGTAGATTTCAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGGA
ATACCACATCCAGCAGGACTAAAAAAGAAAAAGTCAGTAACAGTGCTAGACGTGGGGGACGC
ATATTTTTCAATTCCTTTAGATGAAAGCTTTAGAAAGTATACTGCATTCACCATACCTAGCACA
AACAATGAGACACCAGGAGTAAGGtATCAGTACAACGTGCTTCCACAGGGATGGAAAGGGTC
ACCAGCAATATTTCAGAGTAGCATGACAAAAATCCTAGAGCCTTTTAGACCAAAAAATCCAG
AAATAATCATATATCAATTACATGGATGACTTGTATGTAGGATATGATTTAAAGAAATAGGGC
AGCATAGAGCAAAAATAGAAGAATTATGAGCTCATCTACTGACCTGGGGGCTTCCTACACCA
GACAAAAAGCAT
Sample 51>066B_A1
TGGCAACGACCCATCGTCACAGTAAGGATAGGGGGACAGCTAATAGAAGCTCTCTTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAGATTTGCCAGGAAAATGGACACCAAAAATGA
TAGTGGGAATTGGAGGATTAGTAAAAGTAAAACAGTATGATCAAGTACCTATAGAAATTTGT
GGAAAAAGGGCTGTAGGTGCAGTATTAGTAGGGCCTACACCTGCCAACATCATTGGAAGAAA
CATGTTGACTCAGATTGGTTGTACTCTAAATTTCCCAATTAGCCCTATTGAAACTGTACCAGTA
AAATTAAAGCCAGGAATGGATGGCCCAAGGGTTAAGCAATGGCCATTGACAGAGGAAAAAAT
AAAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAGATTGGGC
CAGAAAATCCATATAATACTCCAATATTTGCTATAAAGAAGAAAAATAGCAATAAATGGAGA
AAATTAGTAGATTTTAGAGAGCTCAATAAAAGAACACAGGACTTTTGGGAAGTTCAATTGGGC
ATACCGCATCCAGCGGGTCTAAARAAGAAAAAATCAGTAACAGTCCTAGATGTGGGTGATGC
ATATTTTTCAGTACCTTTAGATGAAAACTTTAGAAAATACACTGCATTCACCATACCTAGTATA
AACAATGAGACACCAGGAATCAGATATCAATACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCTTTTAGAGCAAAAAATCCAG
AAATAATTATCTATCAATACGTGGATGACTTGTATGTAGCATCTGATTTAGAAATAGGGCAGC
ATAGAGCAAAAATAGAAGAGTTAAGAGCTCATTTATTGAACTGGGGATTTTTTACCCCAGATG
AGAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 52 >8B_D
TGGCAGCGACCCCTAGTCACAGTAAAGATAGGAGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAAATTTACCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGCTTTATCAAAGTAAGACAGTATGATCAAATTCTAGTAGAAATCTGTG
GACATAAAGCTATCGGTACAGTATTAGTAGGACCTACACCTGTCAACATCATTGGAAGAAATT
TGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTACCAGTAAA
ATTAAAGCCAGGGATGGATGGTCCAAAAGTTAAACAATGGCCATTGACAGCAGAAAAAATAA
AGGCACTAACAGAAATTTGTACAGAAATGGAGAAGGAAGGAAAAATTTCAAAAATTGGGCCT
GAAAATCCATATAATACTCCAATATTTGCAATAAAGAAAAAAGACAGTACTGAGTGGAGAAA
GTTAGTAGATTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAATTCAACTAGGAAT
ACCACATCCTGCAGGGTTAGAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCTT
ATTTTTCAGTTCCTTTATATGAAGACTTTAGAAAATATACTGCATTCACTATACCTAGTATAAA
TAATGAGACACCAGGAATTAGATATCAGTACAATGTACTTCCACAAGGATGGAAAGGATCAC
CAGCAATTTTCCAAAGTAGCATGACAAGAATCTTAGAACCTTTTAGAAAACAAAATCCAGAA
ATGATTATTTGTCAATACATAGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAGCAT
AGAATGAAAATAGAGGAATTAAGGGAACATCTGTTGAAGTGGGGATTTACCACACCAGACAA
AAAGCATCAGAAAGAACCTCCATTTCGTTGGATGGGTTATGAACTCCATCCTGATAAATGG
Sample 53 >22_A1
TGGCAACGACCCCTTGTCACAATAAAAATAGGGGGACAAATAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAGATAACTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGGCCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGATTGGTTGCACTCTAAATTTTCCAATTAGTCAGATTGAGACTGTACCAGTGA
AATTGAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATCTGTACAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCAATAAAGAAAAAGGATAGCACTAAATGGAGAA
AAATAGTAGATTTCAGAGAGCTCAATAAAAGAACACAAGACTTCTGGGAAGTTCAATTAGGA

ATACCACATCCAGGGGGCCTAAAAAAGAACAAATCAGTGACAGTACTAGATGTGGGGGATGC
ATATTTTTCAGTACCTCTAGATGAAAGCTTTAGAAAATATACTGCATTCACCATACCTAGTCTA
AACAATGAGACACCAGGAGTCAGGTATCAGTACAATGTGCTCCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAGCAAAAAATCCAG
AAATAGTTATCTACCAATACGTGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGC
ATAGAGCAAAAATAGAGGAGTTGAGAGCTCATCTTTCGAGCTGGGGATTTACTACCCCAGAC
AAAAAGCATCAGAAAGAACATCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 54 >Seq7 _49b_A1D
TGGCAGCGACCCCTCGTTACAGTAAAGATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAAATTTACCAGGAAGATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGCTTTATCAAAGTAAAACAGTATGATCAAATACCAGTAGAAATCTGT
GGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAA
TTTGTTGACTCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTACCAGTA
AAATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAAT
AAAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGACAAATTTCAAAAATTGGGC
CTGAAAATCCATACAATACTCCAATATTTGCAATAAAGAGAAAAGATGGCACTAAATGGAGG
AAATTAGTAGATTTCAGAGAACTTAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGG
AATACCTCATCCTGGAGGGTTAGAAAAGAGAAAATCAGTAACAGTACTAGATGTGGGTGATG
CATATTTTTCAGTTCCCTTACATGAAGACTTTAGAAAGTATACAGCATTCACCATACCTAGTAC
AAACAATGAGACACCAGGAGTTAGATATCAATACAATGTGCTTCCACAAGGATGGAAAGGAT
CACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAACCTTTTAGACAGCAAAATCCAG
AAATGGTTATCTGTCAATACATGGATGATTTGTATGTAGCATCTGATTTAGAAATAGGGCAAC
ATAGAGCAAAAATAGAGGAGTTGAGAGCTCATTTATTGAAGTGGGGACTTACCACACCAGAC
AAAAAGCATCAGAAAGAACCTCCTTTCCGTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 55 >A2_A1
TGGCAACGACCTCTTGTCACAGTAAGAATAGGGGGACAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTAATAGAAGACATAGATTTGCCAGGAAAATGGAAACCAAGAATA
ATAGGGGGAATTGGGGGTCTCATCAAAGTAAGACAATATGATGGGATACCTATAGAAATTTG
TGGAAAAAAGGTTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAA
ACACGTTGACCCAGATTGGATGTACTTTAAATTTCCCAATTAGTCCTATTGACACTGTACCAGT
AAAATTAAAGCCAGGAATGGATGGCCCAAGGGTTAAACAATGGCCATTGACAGAAGAAAAA
ATAAAAGCATTAATAGAAATTTGTACAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGG
GCCTGAAAATCCATACAATACTCCAATATTTGCAATAAAGAAAAAGGATGGCACTAAATGGA
GGAAATTAGTAGATTTCAGAGAGCTTAATAAAAGGACACAAGACTTTTGGGAAGTTCAATTA
GGAATACCGCATCCATCGGGTCTACAAAAGAACAAATCAGTAACAGTACTAGATGTGGGGGA
CGCCTATTTTTCAGTTCCTTTAGATGAAAACTTTAGAAAATATACTGCATTCACCATCCCCAGT
CTAAACAATGCGACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGG
CTCACCAGCAATATTCCAGAGTAGCATGACAAAAATCCTAGAGCCCTTTAGAGCAAAAAACC
CAGAAATAATTATCTATCAATACGTGGATGACTTGTATGTAGGGTCTGATTTAGAAATAGGGC
AGCATAGAGCAAAAATAGAAGAGTTGAGAGCTCACCTATTGAGCTGGGGATTGTTTACACCA
GACAAAAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAG
TGG
Sample 56 >A7_A1C
TGGCAACGACCAGTAGTCACAATAAGAGTAGAAGGACAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAAATTTGTCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAATATGAGCAAGTCCTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTACTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTAAGACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAGGGTCAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGCAATTTGTGAAGAAATGGAAAAGGAAGGAAAAATTACAAAAATTGGGCC
TGAAAATCCATATAACACTCCAATATTTGCAATAAAAAAGAAGGACAGCAATAAGTGGAGAA
AAGTAGTAGATTTCAGGGAACTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGA

ATACCACACCCAGCAGGGTTAAAAAAGAACAAATCAGTGACAGTATTGGATGTGGGGGATGC
ATTTTTTTCAGTTTCTTTAGATGAAGATTTCAGGAAATATACTGCATTTACAATACCAAGTGTA
AACAATGAAGCACCAGGGATTAGATATCAATATAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCGATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTCGAGCAAAAAATCCAG
ACATAGAGATCTGTCAATATGTGGATGACTTGTATGTAGGATCTGACTTAGAAATAGGGCAAC
ATAGAGCAAAAGTAGAAGAGTTAAGAGAACATCTCTTGAGGTGGGGAATTACCACCCCAGAC
AAGAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGATCTCCATCTTGATAAATGG

Sample 57 >A13_A1
TGGCAACGACCCATTGTTACAATAAAGATAGGGGGACAATTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTGTTAGAAGATATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGCGGGAATTGGAGGTTTTGTCAAAGTAAGAGAGTATGATCAAGTACCTATAGAAATCTGTG
GACATAAAGTTATAGGTACAGTATTAGTAGGACCTACACCTTCCAACATAATTGGGAGGAATT
TGTTGACTCAGATTGGCTGCACTTTAAATTTTCCAATTAGTCCTATTGAAACTGTACCAGTAAA
GTTAAAACCAGGGATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAGAAAATAA
AAGCACTAACAGAAATTTGTACAGAGCTGGAAAAGGAAGGAAAAATTTCCAGAATTGGGCCT
GAAAATCCATACAATACTCCAATATTCGCCATAAAGAAAAAGGACAGTACTAAGTGGAGAAA
ATTAGTAGATTTCAGAGAACTTAATAAGAGAACTCAAGATTTCTGGGAAGTTCAACTAGGAAT
ACCACATCCTGCAGGGCTAAAAAAGAAAAAATCAGTAACAGTATTGGACGTGGGTGATGCAT
ATTTTTCAGTTCCCTTAGATGAAGACTTTAGAAAGTATACTGCATTCACCATACCTAGTACAAA
CAATGAGACACCAGGGATTCGATATCAGTACAATGTGCTTCCACAAGGATGGAAAGGATCAC
CAGCAATATTCCAAAGCAGCATGACAAAAATCTTAGAACCTTTTAGAAAACAAAATCCAGGA
ATAATTATCTATCAATACGTGGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGAAACAT
AGAGAAAAAATAGAGGAACTAAGGGAACATCTATTGAAGTGGGGATTTTACACACCAGACAA
AAAACATCAGAAAGAACCTCCATTTCTTTGGATGGGTTATGAACTCCATCCTGATAAATGG
Sample 58 >A14_A1
TGGCAACGACCTATTGTCACAGTAAGAATAGAGGGACAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGACACAGTATTAGAAGACATAAATTTGCCAGGGAAATGGAAACCAAGAATGA
TAGGGGGAATTGGAGGTTTTGTCAAAGTAAAACAGTATGATCAGATAAGTATAGAAATTTATG
GAAAAAAGGCTATGGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTGCCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAGAAAATA
AAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATCACAAAAATTGGGC
CTGAAAATCCATACAATACTCCAATATTTGCAATAAAGAAAAAAGACAGCACTAAATGGAGA
AAATTAGTAGATTTTAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGG
AATACCACATCCAGCGGGCCTAAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGATG
CATATTTTTCAGTCCCTTTACATGAAGAGTTTAGAAAGTATACTGCATTCACCATACCTAGTAC
AAACAATGAGACACCAGGAATCAGGTATCAGTATAATGTGCTTCCACAGGGATGGAAAGGCT
CACCAGCAATATTTCAGAGTAGCATGACAAAAATCTTAGAGCCTTTTAGAATAAAAAATCCAG
AAATAATTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAAC
ATAGAGCAAAAGTAGAGGAGTTAAGAGCCCATCTATTGAGCTGGGGATTTACTACACCAGAT
AAAAAGCATCAAAAAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 59 >A15_A1
TGGCAACGACCCCTTGTCACAGTAAGAATAGGGGGACAATTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAGATTTGCCAGGAAAATGGAAACCAAGAATGA
TAGGGGGAATTGGAGGTTTCATCAAAGTAAAACAGTATGATCAGATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTGGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACCCAGATTGGATGTACTTTAAATTTCCCAATTAGTACTATTGATACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAAATTAAACAATGGCCATTGACAGAAGAAAAGATA
AAAGCATTAATGGAAATTTGTGCAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGACAGTACTAAATGGAGAA
AATTAGTAGATTTCAGAGAACTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGAA

TACCACATCCAGCGGGATTAAAAAAGAACAAATCAGTAACAGTACTAGATGTGGGGGACGCA
TATTTTTCAGTTCCCTTAGATGAAGGCTTTAGAAAGTATACTGCTTTCACCATACCTAGTACAA
ACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTCCCACAGGGATGGAAAGGATCA
CCAGCCATATTCCAGAGTAGCATGACAAAGATCTTAGAGCCCTTTAGATCCAAAAATCCAGAA
ATAATTATTTATCAATACATGGATGATTTGTATGTAGGATCTGATTTAGAAATAGGGCAGCAC
AGAACAAAAATAGAAGAATTGAGAGCCCATCTATTGAGCTGGGGATTTACTACACCAGACAA
AAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAACTCCATCCTGACAAGTGG
Sample 60 >A16_D
TGGCAACGACCCCTCGTCACAGTAAAGATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACGGTATTAGAGGAAATAGATTTGCCAGGGAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAAATACAAGTAGACATCTGT
GGGCACAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAA
TTTATTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTCCTATTGAAACTGTACCAGTG
AAATTAAAGCCAGGGATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAAT
AAAAGCACTAACAGAAATTTGTACAGACATGGAAAAGGAAGGAAAAATTTCAAGAATAGGG
CCTGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGACAGTACTAAGTGGAG
AAAATTAGTAGATTTCAGAGAACTTAATAAGAGGACTCAAGACTTCTGGGAAGTTCAATTAGG
GATACCACATCCTGCAGGGCTAAAAAAGAATAAGTCAGTAACAGTATTGGATGTGGGTGATG
CATATTTCTCAGTTCCCTTACATGAAGAATTTAGAAAATATACTGCATTCACCATACCTAGTAC
AAACAATGAGACGCCAGGAATTAGATATCAGTATAATGTGCTTCCACAAGGGTGGAAAGGAT
CACCAGCAATATTCCAATGTAGTATGACAAAAATCTTAGAACCTTTTAGAAAACAAAACCCAG
ACATAGTTATCTATCAATACGTGGATGATTTGTATGTAGGATCAGACTTAGAAATAGGGCAGC
ATAGAATAAAAATAGAGGAATTAAGGGAACACCTATTGAAGTGGGGGTTTACCACGCCAGAC
AAAAAGCATCAGAAAGAACCTCCATTTCTTTGGATGGGTTATGAACTCCATCCTGATAAATGG
Sample 61>A18_A1
TGGCAACGACCTCTTGTCACAGTAAGAATAGGGGGACAGCTAAAAGAGGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAGAATTGCCAGGGAAATGGAAACCAAAAATG
ATAGGGGGAATTGGGGGTTTCATCAAAGTAAAACAGTATGAACAAATACTCATAGAAATTTG
TGGAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAA
ACATGTTGACCCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCCTATTGAAACTGTACCAGT
AACATTAAAGCCAGGAATGGATGGCCCAAGGGTTAAACAATGGCCATTGACAGAAGAAAAAA
TAAAAGCATTAACAGAAATTTGTACCGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGG
CCTGAAAATCCATACAATACTCCAATATTTGCAATAAAGAAAAAGGATAGCACTAAATGGAG
AAAATTAGTAGATTTCAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTGG
GAATACCACACCCAGCAGGGTTAAAAAGGAGTAAATCAGTAACAGTACTAGATGTGGGGGAT
GCATATTTTTCAGTTCCTTTACATGAGGACTTCAGAAAGTATACTGCATTCACCATACCTAGTA
TAAACAATGAGACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGA
TCACCGGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAATCCA
GAAATGGTTATCTATCAATATATGGATGACTTGTATGTAGGGTCTGATTTAGAAATAGGGCAA
CACAGAGCAAAAATAGAAGAGTTGAGAGCTCATCTATTAAAATGGGGGTTTACTACACCAGA
CAAAAAGCATCAAAAGGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTG
G
Sample 62 >A24S _A1G
TGGCAACGACCCATTGTCACAGTAAAAGTAGAGGGACAGCTAACAGAAGCCCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAGATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATAACTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATCATTGGAAGAAAT
ATGTTGACTCAGATTGGTTGTACTTTAAATTTTCCTATTAGTCCTATTAAAACTGTACCAGTGA
AATTGAAGCCAGGAATGGATGGTCCAAGGGTTAAACAATGGCCATTGACAGCAGAAAAAATA
AAAGCACTAACAGAGATTTGTGCAGAAATGGAAAGGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCGTATAATACTCCAATATTTGCAATAAAGAAAAAGAATAGTACTGAATGGAGGA
AGTTAGTAGATTTCCGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAAGTCCAATTAGGA

ATACCGCATCCAGCGGGCTTAGAAAAGAAAAAATCAGTAACAGTGCTGGATGTGGGGGATGC
ATATTTTTCAGTTCCTTTGGATGAAGATTTTAGAAAGTATACTGCATTCACCATACCCAGCACA
AATAATGCGACACCAGGAATCAGATACCAGTACAATGTACTTCCACAGGGGTGGAAAGGGTC
ACCAGCAATATTTCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAAAACGAAATCCAG
AAGTAGTTATCTACCAATACGTGGATGATTTATATGTAAGCTCTGACTTAGAAATAGGACAAC
ATAGAGCAAAAATAGAGGAATTAAGAGAACATCTACTGAAATGGGGATTAACCACACCAGAT
AAAAAGCATCAGAAAGAACCTCCATTCCTCTGGATGGGATATGAACTCCATCCTGACAAATGG
Sample 63>A34_A1
TGGCAGCGACCTCTTGTCACAGTAAAAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAGAATGA
TAGGGGGAATTGGAGGTTTCATTAAGGTAAAACAGTATGATCAGATACCTATAGAAATTTGTG
GAAAAAAGGCTATAGGCACAGTATTGGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATATTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTGCCAGTAA
AATTAAAGCCAGGAATGGATGGTCCAAAGGTTAAACAATGGCCATTGACAGAGGAGAAAATA
AAAGCATTAACAGAAATCTGTTTAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGACAGCACTAAATGGAGGA
AATTAGTAGATTTCAGAGAGCTCAATAAAAGAACTCAAGATTTTTGGGAAGTTCAATTAGGAA
TACCGCATCCAGCAGGCCTAAAAAAGAACAAATCAGTAACAGTACTAGATGTGGGGGATGCA
TATTTTTCAGTGCCTTTAGATGAAGACTTTAGAAAGTATACTGCGTTCACCATACCTAGTACAA
ACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCGCAGGGATGGAAAGGATCA
CCGGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATTAAAAAATCCAGAG
ATAATTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGCAT
AGAGCAAAAATAGAAGAGTTGAGAGCTCATCTATTGAGCTGGGGATTTACTACCCCAGACAA
AAAACATCAGAAAGAACCACCATTTCTTTGGATGGGATATGAACTCCATCCTGACAAGTGG
Sample 64 >A41B_A1
TGGCAACGACCTCTTGTCACAGTAAAAATAGGGGGACAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTGATCAAAGTAAGACAGTATGAGCAAGTACTTATAGAAATTTGT
GGAAAAAAGGCTATAGGTACAGTATTGGTAGGACCTACTCCTGTCAACATAATTGGAAGAAA
TATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGATACTGTACCAGTA
AAATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTAACAGAAGAAAAAAT
AAAAGCATTAACAGAAATTTGCCTAGAGATGGAAAAGGAAGGGAAAATTACAAAAATTGGGC
CTGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAGGATAGCACTAAATGGAGA
AAATTAGTAGATTTCAGAGAACTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTGGGA
ATACCGCATCCAGCGGGCTTAAAAAAGAAGAAATCAGTAACAGTACTAGATGTGGGGGATGC
ATATTTTTCAGTTCCCTTAGATGAAAACTTTAGAAAGTATACTGCATTTACAATACCTAGTACA
AACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGCATAGCATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAATCCAGA
AATGATTATCTACCAATACATGGATGACTTGTATGTAGGATCTGATTTGGAAATAGGACAACA
TAGAACAAAAATAGAAGAGTTGAGAGCTCATCTATTGAGCTGGGGATTTACTACACCAGACA
AAAAGCATCAGAAGGAACCTCCATTTCTTTGGATGGGATATGAACTTCATCCAGACAAGTGG
Sample 65>A44_A1
TGGCAACGACCCCTCGTCACAATAAAGATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGGTTTATCAAAGTAAAACAGTATGATCAGATACTCATAGAAATTTGTG
GAAAAAAGGCTATAGGAACAGTCTTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACTCAGATTGGTTGTACTTTAAATTTTCCAATCAGTCCTATTGAGACTGTACCAGTAA
AATTAAAGCCAGGAACGGATGGCCCAAGGATTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTGCAGAGATGGAAAAGGAAGGAAAGATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAGAATAGCGATACATGGAGAA
AATTAGTAGATTTTAGAGAGCTCAATAAAAGAACTCAGGACTTCTGGGAAGTTCAATTAGGAA
TACCACATCCCGCAGGTTTAAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGATGCA

TATTTTTCAGTTCCTTTGGATGAAAACTTTAGAAAGTATACTGCATTCACCATACCTAGTATAA
ACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCA
CCAGCAATATTCCAGAGTACCATGACAAAAATCTTAGAGCCCTTCAGGTCACAAAATCCAGAA
ATAGTTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAACAT
AGAGCAAAAATAGAAGAGTTGAGAGCTCATCTATTGAGATGGGGATTTACTACACCAGACAA
AAAGCATCAGAAAGAACCTCCATTTCTTTGGATGGGATATGAACTTCATCCTGACAAATGG
Sample 66>A46_D
TGGCAACGACCCCTCGTCACAGTAAAGATAGGGGGACAGCTAAAGGAAGCTCTATTAGACAC
AGGAGCAGATGATACAGTATTAGAAGAAATGAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGCTTTATCAAAGTAAGACAGTATGAGCAAATACCAGTAGAAATCTGT
GGACATAAAGCTATAGGTACAGTATTAATAGGACCTACACCTGTCAACATAATTGGGAGAAA
TTTGTTGACTCAGATTGGCTGTACTTTAAATTTTCCAATTAGTCCTATTGAAACTGTACCAGTA
AAATTGAAGCCAGGGATGGATGGTCCAAAAGTTAGACAATGGCCATTGACAGAAGAGAAAAT
AAAAGCATTAACAGAAATTTGTCTGGAAATGGAAAAGGAAGGAAAGATTTCAAAAATAGGAC
CTGAAAACCCATACAATACTCCAATATTTGCCATAAAGAGAAAGGACAGTACTAAATGGAGA
AAATTAGTAGATTTCAGAGAACTTAATAAGAGAACTCAAGACTTTTGGGAAGTTCAATTAGGA
ATACCACATCCTGCAGGGCTAAAAAAGAATAAGTCAGTAACAGTACTGGATGTGGGTGATGC
ATTTTTTTCAGTTCCCTTATATGAAGACTTTAGAAAATATACAGCATTCACCATACCTAGTATA
AATAATGAAACACCAGGAATTAGATATCAGTACAATGTGCTCCCACAAGGATGGAAAGGATC
ACCGGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAACCTTTTAGAAAACAAAACCCAG
AAATGTTTATCTGTCAATACATGGATGATTTGTATGTAGGGTCTGACTTAGAAATAGGGCAGC
ATAGAGTAAAAATAGAGGAATTAAGGGAACACCTATTAAAGTGGGGATTTACCACACCAGAC
AAAAAGTATCAAAAAGAACCTCCATTTCTTTGGATGGGTTATGAGCTCCATCCTGATAAATGG
Sample 67>A50_A1
TGGCAACGACCTATTGTCACAGTAAAAATAGAGGGACAGCTAAGAGAAGCTCTGCTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATAAGCATAGAAATTTGT
GGAAAAAAGGCTATAGGTACAGTATTGGTAGGTCCTACCCCTGTCAACATAATTGGAAGAAA
TATGTTGACTCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTAAGACTGTACCAGTA
AAATTAAAGCCAGGGATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAAT
AAAAGCATTAACAGAAATTTGTACAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGC
CTGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGACAGCACTGAATGGAGA
AAGTTAGTAGATTTCAGAGAGTTAAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGA
ATACCGCATCCAGCGGGCTTAAAAAAGAACAAATCAGTAACAGTACTAGATGTGGGAGACGC
ATATTTTTCAGTTCCCCTAGATGAAAACTTTAGAAAGTATACTGCATTCACCATACCTAGTATG
AACAATGAAACACCAGGAATTAGGTATCAGTACAATGTACTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACACAAATCTTAGAGCCCTTTAGATTAAAAAATCCAGA
AATGGTTATTTATCAATACATGGATGACTTGTATGTGGGATCTGATTTAGAAATAGGGCAGCA
TAGAGTAAAAATAGAAGAATTAAGAGATCATCTATTGAAATGGGGATTTACTACACCAGACA
AGAAGCATCAGAAAGAACATCCATTTCTTTGGATGGGATATGAACTCCATCCTGACAAGTGG
Sample 68 >A51_A1
TGGCAACGACCCCTTGTCACAGTAAAAATAGGAGGACAACTAAAAGAAGCTCTACTAGATAC
AGGAGCAGATGATACAGTCTTAGAAGATATAAATTTGCCAGGAAAATGGAAACCAAGAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACCCATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTGTTAGTAGGACCTACACCTGTTAACATAATTGGAAGAAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTACCAGTGA
AATTAAAGCCAGGAATGGATGGCCCAAGAGTTAAACAATGGCCATTGACGGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGACC
TGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAGGACAGCACTAAATGGAGGA
AATTAGTAGATTTCAGAGAGCTCAATAAAAGAACACAAGACTTCTGGGAAGTTCAATTAGGA
ATACCGCATCCAGCGGGTCTAAAAAAGAACAAATCAGTAACAGTACTGGATGTGGGGGACGC
ATATTTTTCAGTTCCTTTGGACGAAAGTTTTAGAAAGTATACTGCGTTCACCATACCTAGTACA

AATAATGCAACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATATTAGAGCCCTTTAGATCAAAAAATCCAG
ACATGATTATCTATCAATACGTGGATGACTTGTATGTAGCATCTGATTTAGAAATAGGGCAGC
ATAGAGCAAAAATAGAAGAATTAAGAGCTCATCTATTGAGCTGGGGACTAACTACACCAGAC
AAAAAACATCAGAAAGAGCCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 69>A54_C
TGGCAGCGACCCCTTGTCTCAATAAAAGTAGGGGGACAAATAAAGGAGGCTCTCTTAGACAC
AGGAGCAGATGATACAGTATTAGAAGAAATAAAATTACCAGGAAATTGGAAACCAAAAATGA
TAGGAGGAATTGGAGGTTTTATCAAAGTAAGACAATATGATCAAATACTTATAGAAATTTGTG
GGAAAAAGGCTATAGGTACAGTACTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACTCAGCTTGGATGCACACTAAATTTTCCAATAAGCCCCATTGAAACTGTACCAGTA
AAATTAAAGCCAGGAATGGATGGCCCAAAGGTCAAACAATGGCCATTGACAGAAGAAAAAAT
AAAAGCATTAACAGCAATTTGTGAAGATATGGAAAAAGAAGGAAAGATTACAAGAATTGGGC
CTGAAAATCCATATAACACTCCAGTATTTGCCATAAAAAAGAAAGACAGTACTAAATGGAGG
AAATTAGTAGATTTCAGGGAGCTTAATAAAAGAACTCAAGATTTTTGGGAAGTTCAATTAGGG
ATACCACACCCAGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGGGATGC
ATATTTTTCAGTTCCTTTAGATGAAGGCTTCAGGAAATATACTGCATTCACCATACCTAGTATA
AACAATGAGACACCAGGAATTAGGTATCAATACAATGTGCTCCCACAAGGGTGGAAAGGGTC
ACCAGCAATATTTCAAAATAGCATGACAAAAATCTTAGAACCTTTTAGAAATCAAAACCCAGA
CATGGTCATCTGTCAATATGTAGATGATTTATATGTAGGATCTGACTTGGAAATAGGGCAACA
TAGAGCAAAAATAGAGGAATTAAGAGAACATCTATTAAGGTGGGGATTTACCACACCAGACA
AAAAATATCAGAAAGAGCCCCCATTTCTGTGGATGGGGTATGAACTCCATCCTGACAAATGG
Sample 70 >A58_A1
TGGCAACGACCCCTTGTCACAATAAAAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTACTAGAAGATATAAATTTGCCAGGGAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAGATACTTATAGAAATCTGTG
GAAAAAAGGCTATAGGTACAGTATTGGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACACAGATTGGCTGTACTTTAAATTTTCCAATTAGTCCTATTGAAACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGTCCAAGGATTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCACTAACAGAAATTTGTAAAGAAATGGAGAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAGTATTTGCTATAAAGAGAAAAGATAGCACTAAATGGAGAA
AATTAGTAGATTTTAGAGAGCTCAATAAAAGAACTCAGGACTTCTGGGAAGTTCAATTAGGAA
TACCGCACCCCGCAGGTTTAAAACAGAACAAATCAGTAACAGTACTGGATGTGGGGGACGCA
TATTTTTCAGTTCCTTTAGATGAAAGCTTTAGAAAGTATACTGCATTCACTATACCTAGTATAA
ACAATGAGACACCAGGAATCAGGTATCAATACAATGTGCTTCCACAGGGCTGGAAAGGATCA
CCGGCAATATTCCAGAGTAGCATGATAAAAATCTTAGAGCCCTTTAGAACAAAAAATCCAGA
AGTAATTATCTATCAATACGTGGATGACTTGTATGTGGGATCTGATTTAGAAATAGGACAGCA
TAGAACAAAAGTAGAGGAATTGAGAGCTCATCTATTGAGTTGGGGATTCACTACCCCAGACA
AAAAGCATCAGAAGGAACCCCCATTTCTTTGGCTGGGATATGAGCTCCATCCCGACAAATGG
Sample 71 >A64_D
TGGCAACGACCCCTTGTTACAATAAAAGTAGAGGGACAGCTAAGGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAAATTTGCAAGGAAAATGGAAACCAAAAATG
ATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGAGCAAGTAGTCATAGAAATCTG
TGGACATAAAGCGGTAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAA
ATTTGTTGACTCAGATTGGCTGCACTTTAAATTTCCCAATTAGTCCTATTAAGACTGTACCAGT
AAAATTAAAGCCAGGAATGGATGGTCCAAGAGTTAAACAATGGCCATTGACAGAAGAAAAAA
TAAAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGG
CCTGAAAATCCATACAATACTCCAATATTTGCCATAAAGAAAAAAGACAGTACTAAGTGGAG
AAAATTAGTAGATTTCAGAGAACTTAATAAAAGAACTCAAGATTTCTGGGAAGTTCAACTAGG
AATACCACATCCTGCAGGGCTAGAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATG
CATATTTTTCAGTTCCCTTAGATGAAGACTTTAGAAAGTATACTGCATTCACCATACCTAGTAT
AAACAATGAGACTCCGGGGATTAGATATCAGTACAATGTACTTCCACAGGGATGGAAAGGAT

CACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAACCCTTTAGAAAGCAAAATCCA
GAAATGGATATCTATCAGTATATGGATGATTTGTATGTAGCATCTGATTTAGAAATAGGGCAG
CATAGAACAAAAATAGAGGAATTAAGGGAACACCTCTTGAAGTGGGGATTTACCACACCAGA
CAAAAAACATCAGAAAGAACCTCCATTTCTTTGGATGGGCTATGAACTCCACCCTGATAAATG
G
Sample 72 >A66_D
TGGCAACGACCCCTCGTCACAGTAAAGATAGGGGGACAGCTAAAGGAAGCTCTATTAGACAC
AGGAGCAGATGATACAGTATTAGAAGAAATGAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGCTTTATCAAAGTAAGACAGTATGAGCAAATACCAGTAGAAATCTGT
GGACATAAAGCTATAGGTACAGTATTAATAGGACCTACACCTGTCAACATAATTGGAAGAAA
TTTGTTGACTCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCCTATTGAAACTGTACCAGTA
AAATTGAAGCCAGGGATGGATGGTCCAAAAGTTAGACAATGGCCATTGACAGAAGAGAAAAT
AAAAGCATTAACAGAAATTTGTCTGGAAATGGAAAAGGAAGGAAAGATTTCAAAAATAGGAC
CTGAAAACCCATACAATACTCCAATATTTGCCATAAAGAGAAAGGACAGCACTAAATGGAGA
AAATTAGTAGATTTCAGAGAACTTAATAAGAGAACTCAAGACTTTTGGGAAGTTCAATTAGGA
ATACCACATCCTGCAGGGCTAAAAAAGAATAAGTCAGTAACAGTACTGGATGTGGGTGATGC
ATTTTTTTCAGTTCCCTTATATGAAGACTTTAGAAAATATACAGCATTCACCATACCTAGTATA
AATAATGAAACACCAGGAATTAGATATCAGTACAATGTGCTCCCACAAGGATGGAAAGGATC
ACCGGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAACCTTTTAGAAAACAAAACCCAG
AAATGTTTATCTGTCAATACATGGATGATTTGTATGTAGGGTCTGACTTAGAAATAGGGCAGC
ATAGAGTAAAAATAGAGGAATTAAGGGAACACCTATTAAAGTGGGGATTTACCACACCAGAC
AAAAAGTATCAAAAAGAACCTCCATTTCTTTGGATGGGTTATGAGCTCCATCCTGATAAATGG
Sample 73 >A68B_A1
TGGCAACGACCCCTTGTCACAATAAGGGTAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGCTTTATCAAAGTAAAACAGTATGATCAGATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGGAAC
ATGTTGACCCAGATTGGATGTACTTTAAATTTCCCAATTAGTCCTATTCAGACTGTACCAGTAA
AATTAAAGCCAGGGATGGATGGCCCAAAGATTAAACAATGGCCATTATCAGAAGAAAAAATA
AAAGCATTGACAGAAATTTGTGGAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAACCCATACAATACTCCAATATTTGCAATAAAGAAAAAAGATAGCACTAAATGGAGAA
AATTAGTAGACTTCAGGGAGCTCAATAAAAGAACACAAGATTTTTGGGAAATTCAATTAGGA
ATACCGCATCCCGCAGGTTTGAAAAAGAAAAAATCAGTATCAGTGCTAGATGTGGGGGACGC
ATATTTTTCAGTTCCTCTAGATGAAGACTTTAGAAAATATACTGCATTCACTATACCTAGTATA
AACAATGAGACACCAGGGATCAGGTATCAGTACAATGTGCTCCCACAGGGATGGAAAGGATC
ACCAGCAATATTTCAGGCTAGCATGACAAAAATCTTAGAGCCCTTTAGAGCAAAAAATCCAG
AAATAATTATCATTCAATACGTGGATGACTTATATGTAGGATCTGATTTAGAAATAGGGCAGC
ATAGAGCAAAAATAGAAGAGTTGAGAGCTCATCTATTGAGTTGGGGACTTACTACACCAGAT
AAAAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAATGG
Sample 74 >A81 _A1C
TGGCAACGACCCCTTGTCACAGTAAAAGTAGGAGGGCAGTTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAAATTTGCCAGGGAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTGGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACCCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCCCATTGAAACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAAGTCAAACAATGGCCATTAACAGAAGAGAAAATA
AAAGCATTAACAGCCATTTGTGAAGATATGGAGAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATATAACACTCCAGTATTTGCCATAAAAAAGAAGGACAGTACTAAGTGGAGAA
AATTAGTAGATTTCAGGGAGCTTAATAAAAGAACTCAAGACTTTTGGGAAGTTCAGTTAGGGA
TACCACACCCAGCAGGGTTAGAAAAGAAAAAGTCAGTGACAGTACTGGATGTAGGGGATGCA
TATTTCTCAGTTCCTTTAGATCCAGACTTTAGAAAGTATACTGCATTTACCATACCTAGTATAA
ACAATGAAACACCAGGGATTAGATATCAATATAATGTGCTTCCACAAGGATGGAAAGGATCA

CCATCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAACAAAAAACCCAGA
AATAGTTATCTATCAATATATGGATGACTTATATGTAGGATCTGACTTAGAGATAGGGCAACA
TAGAGCAAAAATAGAGGAGTTAAGAGAACATCTATTGAAATGGGGATTTACTACACCAGATA
AAAAGCATCAGAAAGAACCTCCATTTCTTTGGATGGGGTATGAACTTCATCCTGACAAATGG
Sample 75 >A82
TGGCAACGACCCCTTGTCACAATAAAGATAGGGGGACAACTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTGTTAGAACATATAGATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGAACAGATACCCATAGAAATTTGT
GGGAAAAGGGCTATAGGTACAGTATTAGTAGGACCCACACCTGTCAACATAATTGGAAGAAA
CATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAAACTGTACCAGTA
AAGTTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAAT
AAAAGCATTAACAGAGATTTGTCTAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGC
CTGAAAATCCATACAATACTCCAATATTTGCAATAAAGAAAAAAGATAGCACTAAATGGAGA
AAAGTAGTAGATTTCAGAGAGCTCAATAAAAGGACACAAGACTTTTGGGAAGTTCAATTAGG
AATACCACATCCAGCAGGCCTAAAAAAGAACAAATCAGTAACAGTACTAGATGTGGGGGACG
CATTTTTTTCAGTTCCTTTATATGAAGAATTTAGAAAATATACTGCATTTACCATACCTAGTAC
AAACAATGAGACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGAT
CCCCGGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATTAAAAAATCCAG
AAATAGTTGTTTATCAATATGTGGATGACTTGTATGTAGCCTCTGATTTAGAAATAGGGCAGC
ATAGAACAAAAATAGAAGAATTGAGAGCTCATCTATTGAGCTGGGGATTTACTACACCAGAC
AAAAAGCATCAGAAAGAACCTCCATTTCTTTGGATGGGATATGAGCTCCATCCTGAAAAGTGG
Sample 76 >A88 _A1C
TGGCAACGACCCCTTGTCACAGTAAGAATAGAAGGGCAGTTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAGATACTAATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAAACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGCAGAAAAAATA
AAAGCATTAACAGAAATTTGTACAGAAATGGAGAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGACAGCAATAAATGGAGGA
AATTAGTAGATTTCAGAGAACTTAATAAAAGAACTCAAGACTTTTGGGAAATTCAATTAGGAA
TTCCGCATCCAGCGGGTCTAAAGAAGAACAAATCAGTAACAATACTAGATGTGGGGGACGCA
TATTTTTCAGTCCCTTTAGATGAAAACTTTAGAAAATATACTGCATTCACCATACCTAGTACAA
ACAATGAGACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCA
CCAGCAATATTCCAGAGTAGTATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAATCCAGAA
ATGGTTATCTATCAATACGTGGATGACTTATATGTAGGATCTGATTTAGAAATAGGGCAACAC
AGAGCAAAAATAGAGGAGTTAAGAGAACATCTATTACAGTGGGGACTGACCACACCAGATAA
GAAACATCAGAAAGAACCCCCATTTCTTTGGCTGGGGTATGAACTCCATCCTGACAAATGG
Sample 77 >A91_A1
TGGCAACGACCCCTTGTTACAATAAAAATAGAGGGACAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGTCAGGAAAATGGAAGCCAAGAATGA
TAGGGGGAATTGGAGGTTTTATCAAGGTAAAACAATATGATCAGATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGCACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACCCAGATTGGCTGTACTCTAAATTTCCCAATTAGTCATATTGAGACTGTGCCAGTAA
AATTAAAACCAGGAATGGATGGCCCAAAGGTTAAACAGTGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTGCAGAAATGGAAAAGGAAGGAAAGATTACAAAAATTGGGCC
TGACAATCCATATAATACTCCAATATTTGCAATAAAGAGAAAAGACAACACTAAATGGAGGA
AATTAGTAGATTTCAGGGAACTCAATAAAAGAACTCAAGATTTTTGGGAAGTTCAATTGGGTA
TACCGCATCCAGCGGGCTTAAAAAAGAAGAAATCAATAACAGTACTAGATGTGGGGGACGCA
TTTTTTTCAATTCCTTTAGATGAAAGCTTTAGACAGTATACTGCGTTTACCATACCTAGTACTA
ACAATGAGACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAGGGATCA
CCAGCAATATTTCAGTATAGCATGACAAAAATCTTAGAGCCCTATAGATTAAAAAATCCAGAC

ATAGTTATCTACCAGTACGTGGATGACTTGTTAGTAGGATCTGATTTAGAAATAGGGCAGCAT
AGAGCAAAAATAGAAGAGCTAAGAGATCATTTATTGAAATGGGGATTAACTACGCCAGACAA
AAAGCATCAGAAAGAACCCCCATTTATTTGGATGGGATATGAACTCCATCCTGACAAATGG
Sample 78 >A98 _A1
TGGCAACGACCCCTTGTCACAGTAAAAATAGGAGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAAATTTGCCAGGAAAATGGAGACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAAATATCTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACTCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCCTATTGAAACTGTACCAGTAA
CATTAAAGCCAGGAATGGATGGCCCAAGGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTGAAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGACC
TGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGATAGCACTAAATGGAGAA
AATTAGTAGATTTTAGAGAGCTCAATAAAAGAACTCAAGACTTTTGGGAAGTCCAATTAGGAA
TACCGCATCCTGCAGGTTTAAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGATGCC
TATTTTTCAGTTCCTTTAGATGAAAATTTTAGAAAGTATACTGCATTCACCATACCTAGTATAA
ACAATGAGACACCAGGAATCAGGTATCAGTATAATGTGCTTCCACAAGGATGGAAAGGATCA
CCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGGTCACAAAATCCAGAA
ATAATTATATATCAATACATGGATGACTTGTATGTAGCATCTGATTTAGAAATAGGGCAGCAT
AGAGCAAAAGTAGAGGAGTTGAGAGCTCATCTATTGAAGTGGGGATTTACTACACCAGATAA
AAAGCATCAGAAAGAGCCTCCCTTTCTTTGGATGGGATATGAACTCCATCCTGACAAATGG
Sample 79 >A99_D
TGGCAACGACCCTTTGTCTCAGTAAAAATAGGAGGACAGCCAAGGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAAATTTGCCAGGAAGATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTGTCAAAGTAAGACAGTATGATCAAGTGCCTATAGAAATTTGTG
GACAAAAGGTTATAGGTACAGTATTAGTAGGACCTACACCTGCCAACATAATTGGAAGAAAT
ATATTATCTCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCCTATCGCAACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGATTAAACAATGGCCATTGACAGAAGAGAAAATA
AAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACCCCGATATTTGCCATAAAGAAAAAAGACAGTACTAAGTGGAGAA
AATTAGTAGATTTCAGAGAACTTAATAAGAGAACTCAAGATTTTTGGGAAGTTCAATTAGGAA
TACCACATCCGGCAGGGCTACACAAGAAAAAATCAGTAACAGTGCTGGATGTGGGTGATGCA
TATTTTTCAGTGCCCTTATATGAAGACTTTAGAAAATATACTGCATTCACCATACCTAGTATAA
ACAATGAGACACCGGGAATTAGATATCAGTACAACGTGCTTCCACAAGGATGGAAAGGATCT
CCGGCAATATTCCAAAGTAGCATGACCAAAATCCTAGAACCTTTCAGAAAACAGAATCCAGA
AATGGTTATCTATCAGTATGTGGATGATTTGTATGTAGCATCTGACTTAGAAATAGGGCAGCA
TAGAGAAAAAATAGAGGAACTAAGGGGGCACCTATTGAAATGGGGACTTACCACACCAGATA
AAAAGCATCAGAAAGAACATCCATTCCTTTGGATGGGTTATGAACTCCATCCTGATAGGTGG
Sample 80 >A101 _A1
TGGCAGCGACCTCTTGTCACAGTAAAAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAGAATGA
TAGGGGGAATTGGAGGTTTCATTAAGGTAAAACAGTATGATCAGATACCTATAGAAATTTGTG
GAAAAAAGGCTATAGGCACAGTATTGGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATATTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTGCCAGTAA
AATTAAAGCCAGGAATGGATGGTCCAAAGGTTAAACAATGGCCATTGACAGAGGAGAAAATA
AAAGCATTAACAGAAATCTGTTTAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGACAGCACTAAATGGAGGA
AATTAGTAGATTTCAGAGAGCTCAATAAAAGAACTCAAGATTTTTGGGAAGTTCAATTAGGAA
TACCGCATCCAGCGGGCCTAAAAAAGAACAAATCAGTAACAGTACTAGATGTGGGGGATGCA
TATTTTTCAGTGCCTTTAGATGTAGACTTTAGAAAGTATACTGCGTTCACCATACCTAGTACAA
ACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCGCAGGGATGGAAAGGATCA
CCGGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATTAAAAAATCCAGAG
ATAATTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGCAT

AGAGCAAAAATAGACGAGTTGAGAACTCATCTATTGAGCTGGGGATTTACTACCCCAGACAA
AAAACATCAGAAAGAACCACCATTTCTTTGGATGGGATATGAACTCCATCCTGACAAGTGG
Sample 81 >A67_A1
TGGCAACGACCCCTTGTCACAATAAGGGTAGGGGGACAGCTRAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGCTTTATCAAAGTAAAACAGTATGATCAGATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGRAGGAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTCAGACTGTACCAGTAA
AATTAAAGCCAGGGATGGATGGCCCAAAGATTAAACAATGGCCATTAACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTGGAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAACCCATACAATACTCCAATATTTGCAATAAAGAAAAAAGATAGCACTAAATGGAGAA
AATTAGTAGATTTCAGGGAGCTCAATAAAAGAACACAAGATTTTTGGGAAATTCAATTAGGA
ATACCGCATCCCGCAGGTTTGAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGACGC
ATATTTTTCAGTTCCTCTAGATGAAGACTTTAGAAAATATACTGCATTCACTATACCTAGTACA
AACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTCCCACAGGGATGGAAAGGATC
ACCAGCAATATTTCAGGCTAGCATGATAAAAATCTTAGAGCCCTTTAGAACAAACAATCCAGA
AATAATTATCTGTCAATACATGGATGACTTATATGTAGGATCTGATTTAGAAATAGGGCAGCA
TAGAGCAAAAATAGAAGAGTTGAGAGCTCATCTATTGAGTTGGGGATTTACTACACCAGATA
AAAAGTATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAATGG
Sample 82 >A52 _D
TGGCAACGACCCCTTGTCACAGTAAAGATAGGGGGACAGCTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAAATATGTGTAGAAATCTGTG
GATATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATT
TATTGACTCAGATTGGTTGCACTCTAAATTTTCCCATTAGTCCTATTGAAACTGTACCAGTAAA
ATTAAAACCAGGGATGGACGGCCCAAAAGTTAAACAGTGGCCGTTGACAGAAGAAAAAATAA
AAGCGCTAGTAGAAATTTGTACAGAAATGGAAAAGGAAGGGAAAATTTCAAGAATTGGGCCT
GAAAATCCATACAATACTCCAATATTTGCCATAAAGAAAAAAGACAGTACTAAGTGGAGAAA
ATTAGTAGATTTCAGAGAACTGAATAAGAGAACGCAAGACTTCTGGGAAGTTCAACTAGGAA
TACCACATCCTGCAGGGCTAAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCA
TATTTTTCAGTTCCCCTAGATGCAAACTTCAGAAAATATACTGCATTCACCATACCTAGTACAA
ACAATGAGACGCCAGGAATTAGATATCAGTACAATGTGCTTCCACAAGGATGGAAAGGATCA
CCGGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAACCCTTTAGAAAACAAAATCCAGA
AATAGTTATCTATCAATACGTGGATGATTTGTATATAGGATCTGACTTAGAATTAGGGCAGCA
TAGAATGAAAATAGAGGAATTAAGGGAACACCTATTGAGGTGGGGATTTACTACACCAGACA
AAAAGCATCAGAAAGAACCTCCGTTTCTTTGGATGGGTTATGAACTCCATCCTGACAAATGG
Sample 83 FASTA >A30S _A1D
TGGCAGCGACCCCTCGTTACAGTAAAGATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAAATTTACCAGGAAGATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGCTTTATCAAAGTAAAACAGTATGATCAAATACCAGTAGAAATCTGT
GGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAA
TTTGTTGACTCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTACCAGTA
AAATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAAT
AAAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGACAAATTTCAAAAATTGGGC
CTGAAAATCCATACAATACTCCAATATTTGTAATAAAGAGAAAAGATGGCACTAAATGGAGG
AAATTAGTAGATTTCAGAGAACTTAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGG
AATACCTCATCCTGGAGGGTTAGAAAAGAGAAAATCAGTAACAGTACTAGATGTGGGTGATG
CATATTTTTCAGTTCCCTTACATGAAGACTTTAGAAAGTATACAGCATTCACCATACCTAGTAC
AAACAATGAGACACCAGGAGTTAGATATCAATACAATGTGCTTCCACAAGGATGGAAAGGAT
CACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAACCTTTTAGACAGCAAAATCCAG
AAATGGTTATCTGTCAATACATGGATGATTTGTATGTAGCATCTGATTTAGAAATAGGGCAAC

ATAGAGCAAAAATAGAGGAGTTGAGAGCTCATTTATTGAAGTGGGGACTTACCACACCAGAC
AAAAAGCATCAGAAAGAACCTCCTTTCCGTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 84 145r >145_A1
TGGCAACGACCCATTGTCACAGTAAAAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGAGGAATTGGGGGTTTCATCAAGGTAAAACAGTATGATCAAATAATGATAGAAATTTGT
GGAAAAAAGGCTATAGGCACAGTGTTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAA
CATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTGCCAGTA
AAATTAAAGCCAGGAATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAGGAAAAAAT
AAAAGCATTAACAGAAATTTGTACAGATATGGAAAAGGAAGGAAAAATTTCAAGAATTGGGC
CTGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGACAGCACTAAATGGAGG
AAATTAGTAGATTTCAGAGAGCTCAATAAAAGAACACAAGATTTCTGGGAAGTTCAATTAGG
AATACCACATCCAGCGGGCTTAAAAAAGAACAAATCAGTAACAGTACTAGATGTGGGGGACG
CATATTTTTCAGTTCCTTTAGATGAAGAGTTTAGAAAATATACTGCATTCACCATACCTAGTAC
AAACAATGAGACACCAGGAACCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGAT
CACCGGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATTAAAAAATCCAG
AGATAATTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGC
ATAGAACAAAAATAGAAGAGTTGAGAGCTCATCTATTGAGCTGGGGATTTACTACACCAGAC
AAAAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 85 >A25_A1
TGGCAACGACCTCTTGTCCCAGTGAAAATAGAGGGACAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGTCAGGAAAATGGAAACCAAGAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTTATAGAAATTTGTG
GRAAAAGGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACTCAAATTGGTTGTACTTTAAATTTCCCAATTAGTCCTGTTGAGACTGTACCAGTAA
AATTAAAGCCAGGGATGGATGGTCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCAATAAAGAAAAAAGATAGCACTAAATGGAGAA
AACTAGTAGATTTCAGAGAACTCAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGGA
ATCCCGCATCCAGCGGGCTTAAAAAAGAAAAAATCAGTAACAGTATTAGATGTGGGGGACGC
ATATTTTTCAGTTCCTTTATATGAAGATTTTAGAAAATATACTGCGTTCACCATACCTAGTATG
AACAATGAGACACCAGGGATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGCATAGTATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAATCCAGA
AATAATTATTTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGAGCAGCA
TAGAGCAAAAATAGAAGAGTTGAGAGCTCATTTATTGCACTGGGGGTTTACTACACCRGACAA
AAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 86 >A39_A1
TGGCAGCGACCCCTTGTTACAGTAAGAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAGATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGGTTCATCAAGGTAAAACAATATGATCAGATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTGGGGAGACCTACACCTGTTCACATAATTGGAAGAAAT
CTGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAAACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCGAGGGTTAAACAATGGCCATTATCAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTACAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAGTATTTGCAATAAAGAAAAAGGATAGCACTAAATGGAGGA
AACTAGTAGATTTCAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGGA
ATCCCGCATCCAGCGGGCTTAAAAAAGAAGAAATCAGTAACAGTGCTAGATGTGGGGGACGC
ATATTTCTCAGTTCCTTTAGATGAAGGTTTTAGAAAATATACTGCATTCACCATACCTAGTACA
AACAATGAAACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATTTTAGAACCCTTTAGAGCAAAGAATCCAG
AAATAATTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGC

ATAGAACAAAAATAGAGGAGTTAAGAGATCATCTATTGAAGTGGGGATTTACTACACCAGAC
AAAAAGCATCAGAAAGAACCCCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 87 >A78_A1
TGGCAACGACCCCTTGTCACAGTAAAAATAGGGGGACAGCTAAAGGAAGCTCTATTAGACAC
AGGAGCAGATGATACAGTATTTGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTCATCAAGGTAAAACAGTATGATCAGATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCAGTCAACATAATTGGAAGAAAT
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAAACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTATGGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGACAGCACTAAATGGAGGA
AACTAGTAGATTTTAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGGA
ATACCGCATCCAGCGGGCCTAAAAAAGAAGAAATCAGTAACAGTACTAGATGTGGGGGACGC
ATATTTTTCAGTTCCTTTAGATGAAAGCTTTAGAAAATATACTGCATTCACCATACCTAGTACA
AACAATGAGACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGGTC
ACCGGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAATCCAG
AAATAATTATTTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGC
ATAGAGCAAAAATAGAAGAATTGAGAGCTCATCTATTGAGCTGGGGGTTTACTACACCAGAT
AAAAAACATCAGAAAGAACCTCCATTTCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 88 >A89_A1
TGGCAACGACCTCTTGTTGCAGTAAAAGTAGCAGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAGATTTACCAGGAAAATGGAAACCAAGAATGA
TAGGAGGAATTGGAGGTTTTATCAAGGTAAAACAGTATGATCAGATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACCCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTACCAGTAA
CATTAAAACCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCACTAACAGAAATTTGTACAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAGTATTTGCTATAAAGAAAAAAGACAGTACTAAATGGAGGA
AATTAGTAGATTTCAGAGAACTTAATAAAAGAACTCAGGACTTCTGGGAAGTCCAATTAGGA
ATACCACATCCAGCAGGTTTAAAAAAGAAAAAGTCAGTAACAGTACTAGATGTAGGGGACGC
ATATTTTTCAGTTCCTTTAGATGAAGGCTTTAGGAAGTATACTGCGTTCACCATACCTAGTATA
AACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTGCCACAGGGGTGGAAAGGATC
ACCGGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAACCCAG
AAATAATTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGACAGC
ATAGAACAAAAATAGAGGAGTTGAGAGCTCATCTATTGAGCTGGGGGTTTACGACACCAGAT
AAAAAGCATCAGAAAGAACCTCCATTTCTTTGGATGGGATATGAACTACATCCTGACAAATGG
Sample 89 >A94_A1
TGGCAACGACCCCTCGTCACAGTAAAAGTAGGAGGACAGATAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAGATTTGCCAGGAAGATGGAAACCAAAAATGA
TAGGGGGAATTGGCGGTTTTATCAAAGTAAAACAGTATGATCAGATAACTATGGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGGAAT
ATGTTGACTCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCATATTGAGACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTAAAGAGATGGAAGAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAGTATTTGCTATAAAGAAAAAAGATAGCACTAAATGGAGAA
AGCTAGTAGATTTTAGAGAGCTCAAYAAAAGAACTCAGGATTTCTGGGAAGTTCAATTAGGA
ATACCACACCCTGCAGGGTTAAAGAAGAAGAAATCAGTAACAGTACTAGATGTGGGAGACGC
ATATTTTTCAGTTCCTTTACATGAAAGCTTTAGAAAATATACTGCGTTCACCATACCTAGTATA
AACAATGAAACACCAGGAATCAGATACCAGTACAATGTGCTTCCACAAGGATGGAAAGGGTC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTCAGATCAAAAAATCCAG
AGATAATTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGC

ATAGAGCAAAAATAGAAGAGTTAAGAGCTCACCTGTTGAGCTGGGGGTTTACTACACCAGAC
AAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAATGG
Sample 90 A27S >A27_A1G
TGGCAACGACCCCTTGTCACAGTAAGAATAGGAGGACAGCTAAAAGAAGCTCTCTTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAAATTTGCCTGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAGATACTTATAGAAATTTGTG
GGAAAAAGGCTATAGGCACAGTGTTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTACCAGTAA
CATTAAAGCCAGGAATGGATGGCCCAAGGATTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAGTAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGATAGCACTAAATGGAGAA
AATTAGTAGATTTTAGAGAGCTCAATAAAAGAACTCAGGATTTCTGGGAAGTTCAATTAGGAA
TACCACATCCCGCAGGTTTAAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGATGCA
TATTTTTCAGTTCCTTTAGATGAAAGCTTTAGAAAGTATACTGCATTCACCATACCTAGTATAA
ACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCG
CCAGCAATATTCCAGAGCAGCATGACAAAAATCTTAGAGCCCTTTAGGTCACAAAACCCAGA
AATAATTATCTATCAATACATGGATGACTTATATGTAGGATCTGATTTAGAAATAGGGCAACA
TAGAGCAAAAGTAGAGGAGTTGAGAGCTCATTTACTGAGTTGGGGATTTACTACACCAGATA
AAAAGCATCAGAAAGAACCTCCATTTCTTTGGATGGGATATGAACTCCATCCTGACAAATGG
Sample 91 A56B >A56_A1
TGGCAGCGACCCCTTGTCACAGTGAAAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCCGATGATACAGTATTAGAAGACATAGATTTGCCAGGAAAATGGAAGCCAAAAATGA
TAGGGGGAATTGGAGGTTTCATCAAGGTAAAACAGTATGATCAGATACTTATGGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAAACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAAATTAAACAATGGCCATTGACAGAGGAAAAAATA
AAAGCATTAACAGAGATTTGTACAGAAATGGAAAAGGAAGGAAAAATTACAAAAATCGGGC
CTGAAAATCCATACAATACTCCCATATTTGCAATAAAGAAAAAAGATAGCACTAAATGGAGA
AAATTGGTAGATTTCAGAGAGCTTAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGG
AATACCGCATCCAGCGGGCCTAAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGACG
CATATTTTTCAGTTCCTTTAGATGAAAGCTTTAGAAAGTATACTGCATTCACCATACCTAGTAT
AAACAATGAGACGCCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGAT
CACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGATCCTTTTAGAGCAAAGAATCCAG
AAATAGTTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGAGCAGC
ATAGAGCAAAAATAGAAGAGTTAAGAGCTCATCTATTGAGCTGGGGATTTACTACCCCAGAC
AAGAAGCATCAGAAAGAACCCCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 92 A19S >A19_A1
TGGCAACGACCCCTTGTCACAGTAAGAATAGGAGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTAGTAGAAGATATAAATTTGCCAGGAAAATGGAAACCAAGAATGA
TAGGGGGCATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAGATACATATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTACTAGTAGGACCTACACCTGTTAACATAATTGGAAGAAAT
ATGTTAACTCAGATTGGCTGTACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTACCAGTAA
CACTAAAGCCAGGAATGGATGGCCCAAGGATTAAACAATGGCCTTTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTAGTGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGACC
TGAAAATCCATACAATACTCCAGTATTTGCTATAAAGAAAAAAGATAGCACTAAATGGAGAA
AATTAGTAGATTTCAGGGAGCTCAATAAAAGAACTCAGGACTTCTGGGAAGTTCAATTAGGA
ATACCACATCCCGCAGGTTTGAAAAAGAAAAAATCAGTGACGGTACTGGATGTGGGGGATGC
ATATTTCTCAGTTCCTTTAGATATAAACTTTAGAAAGTATACTGCATTCACCATACCTAGTATA
AATAATGAGACACCAGGAATAAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAGCAAAAAATCCAG
AAGTAATTATCTATCAGTACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGC

ATAGAACAAAAGTAGAGGAGTTGAGAGCTCATCTATTGAGTTGGGGATTCACTACACCAGAC
AAAAAGCATCAGAAAGAACCTCCATTTCTTTGGATGGGTTATGAACTCCATCCTGACAAATGG
Sample 93 A49B >A49_A1
TGGCAGCGACCTGTTGTCACAGTAAAAGTAGGGGGACAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGACACAGTATTAGAAGACATAGATTTACCAGGGAAATGGAAACCAAGAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAGATAATTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCCACACCTGTCAACATAATTGGAAGGAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTACCAGTAA
AATTAAAACCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAGGAAAAAATA
AAAGCATTAACAGAAATTTGTATAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAACCCATACAATACTCCAATATTTGCAATTAAGAAAAAAGATAGCACCAAATGGAGGA
AATTAGTAGATTTCAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGGA
ATACCACACCCAGCAGGCCTAAAAAAGAAGAAATCAGTAACAGTACTAGATGTGGGGGACGC
ATATTTTTCAGTTCCTTTAGATAAAGACTTTAGGAAATATACTGCATTTACCATACCTAGTACA
AACAATGAGACACCAGGAATCAGATACCAGTACAATGTGCTTCCGCAGGGATGGAAAGGGTC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCCTAGAGCCCTTTAGATCAAAAAATCCAG
AAATAATTATTTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGC
ATAGAACAAAAATAGAAGAGTTAAGAGCTCATCTATTGAGCTGGGGATTTACTACCCCAGAC
AAAAAGCATCAGAAAGAACCTCCATTTCTTTGGATGGGATATGAGCTCCATCCTGACAAATGG
Sample 94 >058J_A1
TGGCAACGACCTATTGTCACAATAAGAATAGGGGGACAGTTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAGAGTTGCCAGGAAAATGGAAACCAAAAATG
ATAGGGGGAATTGGAGGTTTCATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATTTAT
GGAAAAAAGGCTATAGGTACAGTACTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAA
TATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTACCAGTA
AAATTAAAGCCAGGAATGGATGGGCCAAAGGTTAAACAGTGGCCATTGACAGAAGAAAAAAT
AAAAGCATTAACAGAAATTTGTACAGATATGGAGAAGGAAGGAAAAATTTCAAGAATTGGGC
CTGAAAATCCATATAATACTCCAATATTTGCAATAAAGAAAAAAGATAGCACTAAATGGAGG
AAATTAGTAGATTTTAGAGAGCTCAATAAAAGAACACAGGACTTTTGGGAAGTCCAATTAGG
GATACCGCATCCAGCGGGCCTAAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGACG
CATATTTTTCAGTTCCTTTAGATGAGACCTTTAGAAAATATACTGCGTTCACCATACCTAGTAT
AAATAATGAAACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGAT
CACCTGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAGCAAAGAATCCAG
AAATAGTTATCTATCAGTACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAAC
ATAGAGCAAAGATAGAAGAGCTGAGAGTTCATTTATTGAAGTGGGGATTTACTACCCCAGAC
AAAAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 95 >012B_A1
TGGCAACGACCTGTAGTCACAGTAAGAATAGAGGGACAACTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTGGTGGGACCTACACCTGTCAACATAATTGGAAGGAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCCATTGATACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGCACAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGACAGCACTAAATGGAGAA
AATTAGTAGATTTTAGAGAGCTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGGA
TACCGCATCCAGCGGGTTTGAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGACGCC
TATTTTTCAGTTCCCTTAGATGAAGACTTTAGAAAATATACAGCATTCACCATACCTAGTACAA
ACAATGAGACACCAGGAATAAGGTATCAGTACAATGTACTTCCACAGGGATGGAAAGGATCA
CCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAGCAAAAAATCCAGA
AATGGTAATCTATCAGTACATGGATGACTTGTATGTAGGCTCTGATTTAGAAATAGGGCAGCA

TAGAGCAAAAGTAGAAGAATTAAGAAAACATCTATTGAGCTGGGGATTTACTACACCAGACA
AAAAGCATCAGAAAGAACCTCCATTTCTCTGGATGGGGTATGAACTCCATCCTGATAAGTGG
Sample 96 >A41_A1
TGGCAACGACCTCTTGTCACAGTAAAAATAGGGGGACAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTGATCAAAGTAAGACAGTATGAGCAAGTACTTATAGAAATTTGT
GGAAAAAAGGCTATAGGTACAGTATTGGTAGGACCTACTCCTGTCAACATAATTGGAAGAAA
TATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGATACTGTACCAGTA
AAATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTAACAGAAGAAAAAAT
AAAAGCATTAACAGAAATTTGCCTAGAGATGGAAAAGGAAGGGAAAATTACAAAAATTGGGC
CTGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAGGATAGCACTAAATGGAGA
AAATTAGTAGATTTCAGAGAACTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTGGGA
ATACCGCATCCAGCGGGCTTAAAAAAGAAGAAATCAGTAACAGTACTAGATGTGGGGGATGC
ATATTTTTCAGTTCCCTTAGATGAAAACTTTAGAAAGTATACTGCATTTACAATACCTAGTACA
AACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGCATAGCATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAATCCAGA
AATGATTATCTACCAATACATGGATGACTTGTATGTAGGATCTGATTTGGAAATAGGACAACA
TAGAACAAAAATAGAAGAGTTGAGAGCTCATCTATTGAGCTGGGGATTTACTACACCAGACA
AAAAGCATCAGAAGGAACCTCCATTTCTTTGGATGGGATATGAACTTCATCCAGACAAGTGG
Sample 97 >033B_A1
TGGCAACGACCCCTTGTCACAGTAAAAATAGGGGGGCAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAGATTTGCCAGGGAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGGTTCATCAAGGTAAGACAGTATGATCAAATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACTCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGCGACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTACAGATATGGAAAAAGAGGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAGTATTTGCAATAAAGAAAAAGGATAGCACTAAATGGAGAA
AATTAGTAGATTTCAGAGAGCTCAATAAAAGAACACAAGATTTTTGGGAAGTTCAATTAGGG
ATACCGCATCCAGCGGGCCTAAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGACGC
ATACTTTTCAGTTCCTTTAGATGTAAGCTTTAGAAAATATACTGCATTCACCATACCTAGTATA
AATAATGAGACACCAGGAATAAGGTATCAGTACAATGTGCTTCCGCAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTCAGAGCAAATAATCCAG
AAATAATTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATCTAGAAATAGGGCAAC
ATAGAGCAAAAATAGAAGAATTAAGAGCACATCTGTTGAGCTGGGGATTTACTACACCAGAC
AAAAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGAACTCCATCCTGATAAATGG
Sample 98 >A56_A1
TGGCAGCGACCCCTTGTCACAGTGAAAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCCGATGATACAGTATTAGAAGACATAGATTTGCCAGGAAAATGGAAGCCAAAAATGA
TAGGGGGAATTGGAGGTTTCATCAAGGTAAAACAGTATGATCAGATACTTATGGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATCGGAAGAAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTCTCCCATTAGTCCTATCGAAACTGTACCAGTAA
AATTAAAGCCAGGAAAGGATGGCCCAAAAATTAAACAATGGCCATTGACAGAGGAAAAAAT
AAAAGCATTAACAGAGATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATCGGGC
CTGAAAATCCATACAATACTCCCATATTTGCAATAAAGAAAAAAGATAGCACTAAATGGAGA
AAATTGGTAGATTTCAGAGAGCTTAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGG
AATACCGCATCCAGCGGGCCTAAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGACG
CATATTTTTCAGTTCCTTTAGATGAAAGCTTTAGAAAGTATACTGCATTCACCATACCTAGTAT
AAACAATGAGACGCCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGAT
CACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGATCCTTTTAGAGCAAAGAATCCAG
AAATAGTTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGAGCAGC

ATAGAGCAAAAATAGAAGAGTTAAGAGCTCATCTATTGAGCTGGGGATTTACTACCCCAGAC
AAGAAGCATCAGAAAGAACCCCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 99 >A19_A1
TGGCAACGACCCCTTGTCACAGTAAGAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTAGTAGAAGATATAAATTTGCCAGGAAAATGGAAACCAAGAATGA
TAGGGGGCATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAGATACATATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTACTAGTAGGACCTACACCTGTTAACATAATTGGAAGAAAT
ATGTTAACTCAGATTGGCTGTACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTACCAGTAA
CACTAAAGCCAGGAATGGATGGCCCAAGGATTAAACAATGGCCCTTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTAGTGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGACC
TGAAAATCCATACAATACTCCAGTATTTGCTATAAAGAAAAAAGATAGCACTAAATGGAGAA
AATTAGTAGATTTCAGGGAGCTCAATAAAAGAACTCAGGACTTCTGGGAAGTTCAATTAGGA
ATACCACATCCCGCAGGTTTAAAAAAGAAAAAATCAGTGACGGTACTGGATGTGGGGGATGC
ATATTTTCCAGTACCTTTAGATATAAACTTTAGAAAGTATACTGCATTCACCATACCTAGTATA
AATAATGAGACACCAGGAATAAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAAACTTAGAGCCCTTTAGATCAAAAAATCCAG
AAGTAATTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGC
ATAGAACAAAAGTAGAGGAGTTGAGAGCTCATCTATTGAGTTGGGGATTCACTACACCAGAC
AAAAAGCATCAGAAAGAACCTCCCTTTCTTTGGATGGGTGATG
Sample 100 >A111B _A1B
TGGCAACGACCTCTTGTCACAGTAAAAATAGGGGACCAACTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAGATTTGCCAGGGAGATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAAATACCTATAGAAATTTGTG
GAAAAAGGGCTATAGGTACAGTGTTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTACCAGTAA
AATTGAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTACAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCCATAAAGAAAAAGGATAGCTCTAAATGGAGAA
AATTAGTAGATTTCAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGGA
ATACCACATCCAGCAGGACTAAAAAAGAAAAAGTCAGTAACAGTGCTAGACGTGGGGGACGC
ATATTTTTCAATTCCTTTAGATGAAAGCTTTAGAAAGTATACTGCATTCACCATACCTAGCACA
AACAATGAGACACCAGGAGTAAGGTATCAGTACAACGTGCTTCCACAGGGATGGAAAGGGTC
ACCAGCAATATTTCAGAGTAGCATGACAAAAATCCTAGAGCCTTTTAGATCAAAAAATCCAGA
AATAATCATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGCA
TAGAGCAAAAATAGAGGAGCTGAGACAACATCTGTTGAGGTGGGGACTTACCACACCAGACA
AAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGAACTCCATCCTGATAAATGG
Sample 101>11B _A1
TGGCAACGACCCCTTGTCACAGTAAAAATAGGGGGACAGCTAATAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAAATCTGCCAGGAAAATGGAAACCAAAAATG
ATAGGGGGAATTGGAGGCTTTATCAAAGTAAGACAGTATGATCAGATACTTATGGAAATTTGT
GGGAAAAAGGCTATAGGTACAGTATTGGTAGGACCTACACCTGTCAACATAATTGGAAGAAA
TATGTTGACCCAGATTGGCTGTACTTTAAATTTCCCCATTAGTCCTATTGATACTGTACCAGTA
AAATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAAT
AAAGGCATTAACAGAAATTTGTATGGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGC
CTGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGATAGCACTAAATGGAGA
AAATTAGTAGATTTCAGAGAACTCAATAAAAGAACTCAAGACTTCTGGGAAGTTCAGTTAGG
AATACCGCATCCAGCGGGTTTGAAAAAGAACAAATCAGTAACAGTACTAGATGTAGGGGATG
CATATTTTTCGGTTCCATTAGATGAAAGCTTTAGAAAGTATACAGCATTCACCATACCCAGTAC
AAACAATGAGACACCAGGAATCAGGTATCAGTATAATGTGCTTCCACAGGGATGGAAAGGAT
CACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAACCCTTTCGAGCAAAAAATCCA
GAAATAACTATCTATCAATACGTGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAG
CATAGAGCAAAAATAGAAGAACTAAGGGCTCATCTATTGAGCTGGGGACTTACTACACCAGA

CAGAAAGCATCAGAAAGAACATCCATTTCTTTGGATGGGATATGAACTCCATCCTGACAAGTG
G
Sample 102>49 _A1D
TGGCAGCGACCCCTCGTTACAGTAAAGATAGGGGGACAGCTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAAATTTACCAGGAAGATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGCTTTATCAAAGTAAAACAGTATGATCAAATACCAGTAGAAATCTGT
GGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAA
TTTGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTACCAGTA
AAATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAAT
AAAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGACAAATTTCAAAAATTGGGC
CTGAAAATCCATACAATACTCCAATATTTGCAATAAAGAGAAAAGATGGCACTAAATGGAGG
AAATTAGTAGATTTCAGAGAACTTAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGG
AATACCTCATCCTGGAGGGTTAGAAAAGAGAAAATCAGTAACAGTACTAGATGTGGGTGATG
CATATTTTTCAGTTCCCTTACATGAAGACTTTAGAAAGTATACAGCATTCACCATACCTAGTAC
AAACAATGAGACACCAGGAGTTAGATATCAATACAATGTGCTTCCACAAGGATGGAAAGGAT
CACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAACCTTTTAGACAGCAAAATCCAG
AAATGGTTATCTGTCAATACATGGATGATTTGTATGTAGCATCTGATTTAGAAATAGGGCAAC
ATAGAGCAAAAATAGAGGAGTTGAGAGCTCATTTATTGAAGTGGGGACTTACCACACCAGAC
AAAAAGCATCAGAAAGAACCTCCTTTCCGTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 103 >13r _A1
TGGCAACGACCCCTCGTCACAATAAAGATAGGGGGACAACTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAGATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAGGTACTTATGGAAATTTGTG
GACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGGAAT
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGATACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTAACAGAAGAAAAAATA
AAGGCATTAACAGAAATTTGTACAGAGATGGAAAAGGAAGGAAAAATTACAAAAATTGGGCC
TGAAAATCCATATAATACTCCAATATTTGCTATAAAGAAAAAGGATAGCACTAAATGGAGAA
AAGTAGTAGATTTCAGAGAACTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGA
ATACCGCATCCAGCGGGCATAAAAAAGAATAAATCAGTAACAGTACTAGATGTGGGGGATGC
ATATTTTTCAGTTCCCTTAGATGAAAACTTTAGAAAGTACACTGCGTTTACCATACCTAGTACA
AACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGTATAGCATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAATCCAGA
AATAATTATCTATCAATACGTGGATGACTTGTATGTAGGATCTGATTCGGAAATAGGGCAGCA
TAGAACAAAAATAGAAGAGTTGAGAGCTCATCTATTGAGCTGGGGACTTACTACACCAGACA
AGAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAACTCCATCCAGACAAGTGG
Sample 104 >107J_A1
TGGCAACGACCTCTTGTCACAGTAAAAGTAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTGTTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTCATCAAGGTAAAACAATATGATCAGGTAATTATGGAAATTTGTG
GGAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGGAAC
ATGTTGACACAAATTGGTTGTACTCTAAATTTTCCAATTAGTCCTATTGCGACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAGGCATTAACAGAAATTTGTTTAGAACTGGAGAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCAATAAAGAAAAAAGATAGAACTAAATGGAGGA
AATTAGTAGACTTCAGAGAGCTTAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGGA
ATACCACATCCAGGGGGCTTAAAAAAGAATAAATCAGTAACAGTGCTAGATGTAGGGGACGC
ATATTTTTCAGTCCCTTTAGATGAAGACTTTAGAAAGTATACTGCCTTCACCATACCTAGTACA
AACAATGAGACACCAGGAGTCAGATATCAGTACAATGTGCTCCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAGCAAAAAATCCAG
AAATAACTATCTATCAATATGTGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGC

ACAGAACAAAAATAGAAGAGCTGAGAGCTCATCTATTGAGCTGGGGATTTTTTACACCAGAC
AAAAAGCATCAGAAAGAGCCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 105>141_A1
TGGCAACGACCTCTTGTCACAGTAAAAATAGAGGGACAGTTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTCATCAAAGTAAAACAGTATGATCAGATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTGGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGATACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGACAGCACTAAATGGAGAA
AATTAGTAGATTTCAGAGAGCTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGAA
TACCACATCCAGCGGGCTTAGAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGACGCA
TATTTTTCAGTTCCCTTAGATGAAAGCTTTAGAAAGTATACAGCATTCACCATACCTAGTACAA
ACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCA
CCAGCAATATTCCAGAGTTCCATGACAAAAATCTTAGAGCCCTATAGATCAAAAAATCCAGAG
ATAATTATCTATCAATACGTGGATGACTTGTATGTAGCATCTGATTTAGAAATAGGGCAGCAT
AGAGCAAAAGTAGAAGAATTAAGAAATCATCTATTGAGCTGGGGATTTACTACCCCAGACAA
AAAGCATCAGAAAGAACCTCCATTTCTTTGGATGGGATATGAACTCCATCCTGACAAGTGG
Sample 106>32r _A1
TGGCAACGACCCTTTGTCAAAATAAAGATAGAGGGGCAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAGAATTGCAAGGAAAATGGAAACCAAAAATG
ATAGGGGGAATTGGAGGCTTCATTAAGGTAAAACAGTATGAACAAGTACTTATAGAAATTTGT
GGAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAA
CATGTTGACCCAGATTGGCTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTACCAGTA
AAATTGAAGCCAGGAATGGATGGCCCAAAGATTAAACAATGGCCATTAACAGAAGAGAAAAT
AAAAGCATTAACAGACATTTGTAATGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGC
CTGAAAATCCATACAATACCCCAATATTTGCAATAAAGAAAAAAGATAGCACTAAATGGAGA
AAATTAGTAGATTTCAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGG
AATACCACATCCAGGGGGCCTAAAAAAGAAAAAATCAGTGACAGTACTGGATGTGGGGGACG
CATATTTTTCAGTCCCCTTAGATAAAGAATTTAGAAAATATACTGCATTCACCATACCTAGTAC
AAATAATGAGACACCAGGAATAAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGAT
CACCAGCAATATTCCAAAGTAGTATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAATCCAG
AAATAGTTATCTGTCAGTACGTGGATGACTTGTATGTAGGATCTGATTTAGAAATAGAGCAGC
ACAGAGCAAAAATAGAAGAATTAAGAGCCCATTTATTGAAATGGGGATTTACTACACCAGAC
AAAAAGCATCAGAAGGAACCTCCATTCCGTTGGATGGGATATGAGCTCCATCCTGAAAAGTG
G
Sample 107>A24_A1G
TGGCAACGACCCCTCGTCACAGTAAAAGTAGAGGGACAGCTAACAGAAGCCCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAGATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATAACTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATCATTGGAAGAAAT
ATGTTGACTCAAATTGGTTGTACTTTAAATTTTCCTATTAGTCCTATTAAAACTGTACCAGTGA
AATTGAAGCCAGGAATGGATGGTCCAAGGGTTAAACAATGGCCATTGACAGCAGAAAAAATA
AAAGCACTAACAGAGATTTGTGCAGAAATGGAAAGGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCGTATAATACTCCAATATTTGCAATAAAGAAAAAGAATAGTACTGAATGGAGGA
AGTTAGTAGATTTCCGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAAGTCCAATTAGGA
ATACCGCATCCAGCGGGCTTAGAAAAGAAAAAATCAGTAACAGTGCTGGATGTGGGGGATGC
ATATTTTTCAGTTCCTTTGGATGAAGATTTTAGAAAGTATACTGCATTCACCATACCCAGCACA
AATAATGCGACACCAGGAATCAGATACCAGTACAATGTACTTCCACAGGGGTGGAAAGGGTC
ACCAGCAATATTTCAGAGTAACATGACAAAAATCTTAGAGCCCTTTAGAAAACGAAATCCAG
AAGTAGTTATCTACCAATACGTGGATGATTTATATGTAAGCTCTGACTTAGAAATAGGACAAC

ATAGAGCAAAAATAGAGGAATTAAGAGAACATCTACTGAAATGGGGATTAACCACACCAGAT
AAAAAGCATCAGAAAGAACCTCCATTCCTCTGGATGGGATATGAACTCCATCCTGACAAATGG
Sample 108>A29_A1
TGGCAACGACCCCTCGTCACAATAAAGGTAGGGGGACAACTAAGAGAAGCTCTAATAGATAC
AGGAGCAGATGATACAATAATAGAAGACATAAATTTGCCAGGGAAATGGAAACCAAGAATA
ACAGGGGGACTTGGAGGTTTCATCAGAGTAAAACAGTATGATCAGATAGTTATAGAAATTTGT
GGAAAAAAGGCTATAGGTTCAGTATTAGTAGGACCCACACCAATGAACATAATTGGAAGAAA
CATATTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTACCAGTA
AAACTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGTCAGCAGAAAAAAT
AAAAGCACTAACAGAAATATGTGCAGAAATGGAAAAGGAAGGACAAATTTCAAAAATTGGG
CCTGAAAATCCATACAATACTCCAGTATTTGCAATAAAGAAAAAGAACAGCACTAGATGGAG
AAAATTAGTAGACTTCAGAGAGCTCAATAAGAGAACACAAGATTTTTGGGAAGTTCAATTGG
GAATACCGCATCCAGCGGGCCTAAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGATG
CATTTTTTTCAGTTCCTTTACATGAAGATTTTAGAAAGTATACTGCGTTCACCATACCTAGTAC
AAACAATGAGACACCAGGAATTAGGTATCAGTACAATGTGCTTCCTCAGGGATGGAAAGGAT
CACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAACCCTTTAGATCAAGAAATCCAG
AAATAATCATCTGTCAATACGTGGATGACTTATATGTAGGATCTGATTTGGAAATAGGACAGC
ATAGAGCAAAAATAGAAGAGCTGAGAGCTCATCTGTTGAGCTGGGGGTTTTTTACACCAGACC
AAAAATATCAGAAGGAACCTCCATTCCAATGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 109>A20r_A1B
TGGCAGCGACCTCTTGTCACAGTAAAAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAGAATGA
TAGGGGGAATTGGAGGTTTCATTAAGGTAAAACAGTATGATCAGATACCTATAGAAATTTGTG
GAAAAAAGGCTATAGGCACAGTATTGGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATATTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTGCCAGTAA
AATTAAAGCCAGGAATGGATGGTCCAAAGGTTAAACAATGGCCATTGACAGAGGAGAAAATA
AAAGCATTAACAGAAATCTGTTTAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGACAGCACTAAATGGAGGA
AATTAGTAGATTTCAGAGAGCTCAATAAAAGAACTCAAGATTTTTGGGAAGTTCAATTAGGAA
TACCGCATCCAGCAGGCCTAAAAAAGAACAAATCAGTAACAGTACTAGATGTGGGGGATGCA
TATTTTTCAGTGCCTTTAGATGAAGACTTTAGAAAGTATACTGCGTTCACCATACCTAGTACAA
ACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCGCAGGGATGGAAAGGATCA
CCGGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATTAAAAAATCCAGAG
ATAATTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGCAT
AGAGCAAAAATAGAAGAGTTGAGAGCTCATCTATTGAGCTGGGGATTTACTACCCCAGACAA
AAAACATCAGAAAGAACCACCATTTCTTTGGATGGGATATGAACTCCATCCTGACAAGTGG
Sample 110>A45_A1
TGGCAACGACCCCTCGTCACAATAAAGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAAATCTGCCAGGAAAATGGAAACCAAAAATG
ATAGGGGGAATTGGAGGCTTTATCAAAGTAAGACAGTATGATCAAATACAAGTAGAAATTTG
TGGACAGAGAGCTACAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGGA
ATCTATTGACTCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTACCAGT
AAAATTAAAACCAGCAATGGATGGTCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAA
TAAAAGCATTAACAGAAATTTGTACAGAAATGGAAAAAGAAGGAAAAATTTCAAAAATTGGG
CCTGAAAATCCATATAATACTCCAATATTTGCCATAAAGAAAAAAGACAGTACTAAGTGGAG
AAAATTAGTAGATTTCAGAGAACTTAATAAGAGAACACAAGACTTCTGGGAGGTCCAATTAG
GAATACCTCATCCCGCGGGTTTAAAGAAGAACAAATCAGTAACAATACTAGATGTGGGGGAT
GCATATTTTTCAGTTCCTTTAGATGAAAACTTTAGAAAATATACCGCATTCACCATACCTAGTG
TAAATAATGAGACACCAGGGATTAGATACCAGTACAATGTGCTTCCACAAGGATGGAAAGGA
TCACCAGCAATATTTCAAAGTAGTATGACAAAAATCTTAGAACCTTTTAGAAAACAAAACCCA
GAAATAGTTATCTATCAATACATGGATGATTTATATGTAGGATCTGACTTAGAAATAGGACAG
CATAGAATAAAAATAGAGGAATTAAGGGAACACCTATTGAAGTGGGGATTTACCACACCAGA

CAAAAAGCACCAGAAAGAACCACCATTTCTTTGGATGGGATATGAACTCCATCCTGACAAATG
G
Sample 111>A60_A1
TGGCAGCGACCCCTTGTCACAGTAAAAGTAGGGGGGCAGCTAAAAGAAGCTCTTTTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAGATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAGGTAAAACAGTATGATCAGATACTTATAGAAATTTGTG
GAAAAAGGGCTATAGGTACAGTATTAGTAGGACCTACGCCTGTCAACATAATTGGAAGAAAT
ATGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTGCCAGTAA
AATTAAAGCCAGGAATGGATGGCCCTAGGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTATGGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCAATAAAGAAAAAGGATAGCACTAAATGGAGAA
AATTAATAGATTTCAGAGAACTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAACTAGGA
ATACCGCATCCAGCGGGTTTAAAAAAGAAAAAATCAGTAACAGTTCTAGATGTGGGAGACGC
ATATTTTTCAGTTCCCTTAGATGAAAGTTTTAGAAAGTATACAGCGTTCACCCTGCCTAGTACA
AACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCATTTAGATCAAAAAATCCAG
AGATAATTATCTGCCAATACGTGGATGATTTGTATGTAGGATCTGATTTAGAAATAGAGCAGC
ATAGAGCAAAAATAGAAGAGTTAAGGGCTCATTTATTGAGCTGGGGACTTACTACCCCAGAC
AAAAAGTATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAACTCCATCCTGACAAATGG
Sample 112 >A61r_A1
TGGCAACGACCCCTTGTCACAATAAGGGTAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGCTTTATCAAAGTAAAACAGTATGATCAGATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGGAAC
ATGTTGACCCAGATTGGATGTACTTTAAATTTACCAATTAGTCCTATTCAGACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGATTAAACAATGGCCATTATCAGAAGAAAAAATA
AAAGCATTGACAGAAATTTGTGGAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAACCCATACAATACTCCAATATTTGCAATAAAGAAAAAAGATAGCACTAAATGGAGAA
AATTAGTAGATTTCAGGGAGCTCAATAAAAGAACACAAGATTTTTGGGAAATTCAATTAGGA
ATACCGCATCCCGCAGGTTTGAAAAAGAAAAAATCAGTATCAGTGCTAGATGTGGGGGACGC
ATATTTTTCAGTTCCCTTAGATGAAGACTTCAGGAAGTATACTGCATTTACCATACCTAGTATA
AACAATGGGACACCAGGGATCAGGTATCAGTACAATGTGCTCCCACAGGGATGGAAAGGATC
ACCAGCAATATTTCAAGCTAGCATGACAAAAATCTTAGAGCCCTTTAGAGCAAAAAATCCAG
AAATAATTATCATTCAATACGTGGATGACTTATATGTAGGATCTGATTTAGAAATAGGGCAGC
ATAGAGCAAAAATAGAAGAGTTGAGAGCTCATCTATTGAGTTGGGGACTTACTACACCAGAT
AAAAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGAACTCCATCCTGATAAATGG
Sample 113 >A73_D
TGGCAACGACCCCTCGCTACAATAAAGATAGGGGGGCAGTTAAAAGAAGCTTTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAATATGATCAGATACCTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACTCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTGTTGAGACTGTACCAGTAG
CTTTGAAGCCAGGAATGGACGGCCCAAGGGTTAAACAGTGGCCATTGACAGAAGAAAAATTA
AAAGCCTTAACAGAAATTTGTACAGAGATGGAAAAAGAAGGAAAAATTTCAAGAATAGGGCC
CGAAAATCCATACAATACTCCAGTATTTGCCATAAAGAGAAAAGACGGTACTAAGTGGAGAA
AATTAATAGATTTCAGAGAACTTAATAAAAGAACTCAAGACTTCTGGGAAATTCAACTAGGA
ATACCACATCCGGCAGGGATAAAAAAGAACAAATCAGTAACAGTTCTGGATGTGGGTGATGC
ATATTTTTCAGTTCCCCTGTATGAAGACTTTAGAAAATATACTGCATTCACCATACCTAGTATA
AACAATGAAACACCAGGAATTAGATATCAGTATAATGTGCTTCCACAAGGATGGAAGGGATC
ACCGGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAACCTTTTAGAAAACAAAATCCAG
AAATGGTTATCTATCAATACATAGATGATTTGTATGTAGGGTCTGACTTAGAAATAGGGCAGC

ATAGACTAAAAATAGAGGAGTTAAGGGAACATCTGTTGAAATGGGGATTTACCACGCCAGAC
AAAAAGCATCAGAAAGAACCCCCATTTCTTTGGATGGGGTATGAACTCCATCCTGACAAATGG
Sample 114 >A93_A1
TGGCAGCGACCAGTAGTTACAGTAAGAATAGGGGAACAATTGAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTAATAGAAGAAATAAATTTGCCAGGAAAATGGAGACCAAAAATG
ATAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAATATGATCAGATACAGATAGAAATTTG
TGGGAGGAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGGA
ATATGTTGACTCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTAGAACTGTACCAGT
AAAATTAAAGCCAGGAATGGATGGCCCAAGGGTTAAACAATGGCCATTGACAGAAGAAAAA
ATAAAAGCATTAACAGCAATTTGTGCAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGG
GCCTGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGACAGTACTGAATGGA
GAAAATTAGTAGATTTCAGAGAGCTCAATAAAAGAACTCAAGACTTTTGGGAAGTGCAATTA
GGAATACCACATCCATCGGGCCTAAAGAACAATAAATCAGTAACAGTACTGGATGTGGGGGA
TGCATATTTTTCAGTTCCTTTAGATGAAGACTTTAGAAAGTATACTGCATTCACCATACCTAGT
ACAAACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTGCCACAGGGATGGAAAGG
ATCACCAGCAATATTCCAGTCTAGCATGACAAAAATTTTAGAACCCTTCAGATCAAAAAATCC
AGAGATAATTATTTATCAATACGTGGATGACTTGTATGTAGGATCTGATTTGGAAATAGAGCA
GCATAGAGAAAAAGTAGAAGAGCTGAGAGCTCATCTATTGAAATGGGGGTTTACTACCCCAG
ACAAAAAGCATCAGAAAGAACCACCATTTACTTGGATGGGATATGAACTTCACCCTGATAAAT
GG
Sample 115 >A106_A1
TGGCAACGACCCCTCGTCACAATAAAGATAGCAGGACAGCTAAAGGAAGCTCTTTTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAGATTTGCCAGGGAAATGGAAACCAAAAATGA
TAGGAGGAATTGGAGGTTTCATCAAAGTAAAACAGTATGATCAGATACTTATAGACATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACTCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAAACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGGCCAAAGGTTAAACAATGGCCATTGACAGCAGAAAAAATA
AAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAGATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCAATAAAGAGAAAGGATGGCACTAAATGGAGGA
AATTAGTAGATTTCAGAGAGCTTAATAAAAGAACACAAGACTTTTGGGAAGTTCAGTTAGGG
ATACCGCATCCAGCGGGCATAAAAAAGAACAAATCAGTAACAGTACTAGATGTGGGGGACGC
TTATTTTTCAGTCCCTTTAGATGAAAGCTTTAGAAAGTATACTGCATTCACTATACCTAGTACA
AACAATGAGACACCAGGGATAAGGTATCAGTATAATGTACTGCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAAAGTAGTATGACAAAAATCTTAGAGCCCTTTAGAGCAAAAAACCCAG
AAATAATTATCTATCAGTACGTGGATGACTTGTATGTAGGGTCTGATTTAGAAATAGGGCAAC
ATAGAGCAAAAGTAGAAGAGCTGAGAGCTCACCTATTAAAATGGGGACTTACTACACCAGAC
GAAAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAATGG
Sample 116 >A113_A1
TGGCAACGACCCATTGTCACAATAAAAATAGAGGGACAACTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGTGGTTTTATCAAAGTAAGACAGTATGATCAGATACTTATAGAAATTTGTG
GGAAAAAGGCTATAGGTACAGTATTGGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACCCAGATTGGTTGCACTCTAAATTTTCCAATTAGTCCTATTGAAACTGTACCAGTAA
AATTAAAACCAGGAATGGACGGCCCAAAGGTTAAACAATGGCCCTTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATCTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATAGGGC
CTGAGAATCCATACAATACTCCAGTATTTGCTATAAAGAGAAAAGATAACACTAAATGGAGA
AAATTAGTAGATTTTAGAGAGCTCAATAAGAGAACTCAAGACTTTTGGGAAGTCCAATTAGGA
ATACCACATCCTGCAGGGTTACAAAAGAATAAGTCAGTAACAGTACTAGATGTGGGGGATGC
ATATTTTTCAGTTCCCTTAGATGAAGACTTTAGAAAGTATACTGCATTCACCATACCCAGTATA
AACAATGAAACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAAGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAGCAAAAAATCCAG
AAATAGTTATCTATCAATACGTGGATGACTTGTATGTAGCATCTGATTTAGAAATAGGGCAGC

ATAGAGCAAAGATAGAGGAGTTAAGAGCCCATCTGTTGAGCTGGGGATTTACTACACCAGAC
AAAAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAACTCCATCCTGACAAGTGG
Sample 117 >44r_A1
TGGCAACGACCCCTTGTCACAATAAAAATAGGGGGACAAATAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAGATAACTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGGCCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGATTGGTTGCACTCTAAATTTTCCAATTAGTCAGATTGAGACTGTACCAGTGA
AATTGAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATCTGTACAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCAATAAAGAAAAAGGATAGCACTAAATGGAGAA
AAATAGTAGATTTCAGAGAGCTCAATAAAAGAACACAAGACTTCTGGGAAGTTCAATTAGGA
ATACCACATCCAGGGGGCCTAAAAAAGAACAAATCAGTGACAGTACTAGATGTGGGGGATGC
ATATTTTTCAGTACCTCTAGATGAAAGCTTTAGAAAATATACTGCATTCACCATACCTAGTCTA
AACAATGAGACACCAGGAGTCAGGTATCAGTACAATGTGCTCCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAGCAAAAAATCCAG
AAATAGTTATCTACCAATACGTGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGC
ATAGAGCAAAAATAGAGGAGTTGAGAGCTCATCTTTCGAGCTGGGGATTTACTACCCCAGAC
AAAAAGCATCAGAAAGAACATCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 118 >35_A1C
TGGCAGCGACCCCTTGTCTCAATAAAAGTAGGAGGGCAGATAAAAGAGGCTCTCTTAGACAC
AGGAGCAGATGATACTGTATTAGAAGAAATAAAATTGCCAGGAAATTGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATGAAATACCTATAGAAATCTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTGGGACCAACACCTGTCAACATAATTGGAAGAAAT
CTGTTGACTCAGCTTGGATGCACACTAAATTTTCCAATTAGTCCCATTGAAACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTCAAACAATGGCCATTGACAGAAGAAAAGATA
AAAGCATTAACAGAAATTTGTGCAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATAGGGCC
TGAAAATCCATACAATACTCCAATATTTGCAATAAAGAAAAAGAGTGGTAAATGGAGAAAAT
TAGTAGATTTCAGAGAGCTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGAATAC
CACATCCAGCAGGTTTAGAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGATGCATAT
TTTTCAGTTCCCTTAGATGAAAGTTTTAGAAAATATACAGCATTCACCATACCTAGTACAAAC
AATGCAACACCAGGAGTCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACC
AGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAGCAAAATATCCAGAAA
TAATTATCTATCAATACGTCGATGACTTGTATGTAGCATCTGATTTAGAAATAGGGCAGCATA
GAGCAAAAGTAGAAGAATTAAGAGATCATCTGTTGAAATGGGGATTTTTTACACCAGACCAA
AAGCATCAGAAAGAACCTCCATTTCTTTGGATGGGATATGAACTCCATCCTGACACATGG
Sample 119 >49B _A1D
TGGCAGCGACCCCTCGTTACAGTAAAGATAGGGGGACAGCTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAAATTTACCAGGAAGATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGCTTTATCAAAGTAAAACAGTATGATCAAATACCAGTAGAAATCTGT
GGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAA
TTTGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTACCAGTA
AAATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAAT
AAAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGACAAATTTCAAAAATTGGGC
CTGAAAATCCATACAATACTCCAATATTTGCAATAAAGAGAAAAGATGGCACTAAATGGAGG
AAATTAGTAGATTTCAGAGAACTTAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGG
AATACCTCATCCTGGAGGGTTAGAAAAGAGAAAATCAGTAACAGTACTAGATGTGGGTGATG
CATATTTTTCAGTTCCCTTACATGAAGACTTTAGAAAGTATACAGCATTCACCATACCTAGTAC
AAACAATGAGACACCAGGAGTTAGATATCAATACAATGTGCTTCCACAAGGATGGAAAGGAT
CACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAACCTTTTAGACAGCAAAATCCAG
AAATGGTTATCTGTCAATACATGGATGATTTGTATGTAGCATCTGATTTAGAAATAGGGCAAC

ATAGAGCAAAAATAGAGGAGTTGAGAGCTCATTTATTGAAGTGGGGACTTACCACACCAGAC
AAAAAGCATCAGAAAGAACCTCCTTTCCGTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 120 >100r_A1
TGGCAACGACCCCTCGTCACAATAAAGATAGGGGGACAGATAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAGGTACTTATGGAAATTTGTG
GAAAAAGAGCTACAGGTACAGTATTAGTAGGACGTACACCTGTCAACATAATTGGAAGGAAT
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGATACTGTACCAGTAA
ACTTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTAACAGAAGAAAAAATA
AAGGCATTAACAGAAATTTGTACAGAGATGGAAAAGGAAGGAAAAATTACAAAAATTGGGCC
TGAAAATCCATATAATACTCCAATATTTGCTATAAAGAAAAAGGATAGCACTAAATGGAGAA
AAGTAGTAGATTTCAGAGAACTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGA
ATACCGCATCCAGCGGGCATAAAAAAGAATAAATCAGTAACAGTACTAGATGTGGGGGATGC
ATATTTTTCAGTTCCCTTAGATGAAAACTTTAGAAAGTACACTGCGTTTACCATACCTAGTACA
AACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAAGGATGGAAAGGATC
ACCAGCAATATTCCAGTATAGCATGACAAAGATCTTAGAGCCCTTTAGATCAAAAAATCCAGA
AATAACTATCTATCAATACGTGGATGACTTGTATGTAGGATCTGATTCGGAAATAGGACAACA
TAGAACAAAAATAGAAGAGTTAAGAGCTCATCTATTGAGCTGGGGACTTACTACACCAGACA
AGAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAACTCCATCCAGACAAGTGG
Sample 121 >142 _A1
TGGCAACGACCCCTTGTTACAGTAAAAATAGGGGGACAATTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTTGAAGAGATAGATTTACCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTCATCAAAGTAAAACAGTATGATCAGATACTTATAGAAATTTGTG
GAAAAAGGGCTATAGGTACAGTATTAGTAGGACCAACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTACTACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATT
AAAGCGTTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAACCCATACAATACTCCAGTATTTGTTATAAAGAGAAAAGACAACACTAAATGGAGAA
AATTAGTAGATTTCAGAGAACTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGAA
TACCGCACCCAGCGGGCTTAAAAAAGAATAAATCAGTAACAGTACTAGATGTGGGGGACGCA
TATTTTTCAGTTCCCTTAGATAAAGACTTTAGAAAGTATACTGCATTTACCATACCTAGTCTAA
ACAATGAAACACCAGGAATCAGGTATCAGTACAATGTACTTCCACAGGGATGGAAAGGATCA
CCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATTAAAAAATCCAGAA
ATAATTATCTATCAATACGTGGATGACTTGTATGTAGGATCTGATTTGGAGATAGGGGAGCAT
AGAACAAAAATAGAAGAGTTGAGAGCTCATCTATTGAGCTGGGGATTTACTACACCAGACAA
AAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGATAAGTGG
Sample 122 >A6r _A1
TGGCAACGACCAGTAGTCACAATAAGAGTAGAAGGACAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAAATTTGTCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAATATGAGCAAGTCCTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTACTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTAAGACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAGGGTCAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGCAATTTGTGAAGAAATGGAAAAGGAAGGAAAAATTACAAAAATTGGGCC
TGAAAATCCATATAACACTCCAATATTTGCAATAAAAAAGAAGGACAGCAATAAGTGGAGAA
AAGTAGTAGATTTCAGGGAACTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGA
ATACCACACCCAGCAGGGTTAAAAAAGAACAAATCAGTGACAGTATTGGATGTGGGGGATGC
ATTTTTTTCAGTTTCTTTAGATGAAGATTTCAGGAAATATACTGCATTTACAATACCAAGTGTA
AACAATGAAGCACCAGGGATTAGATATCAATATAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCGATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTCGAGCAAAAAATCCAG
ACATAGAGATCTGTCAATATGTGGATGACTTGTATGTAGGATCTGACTTAGAAATAGGGCAAC

ATAGAGCAAAAGTAGAAGAGTTAAGAGAACATCTCTTGAGGTGGGGAATTACCACCCCAGAC
AAGAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGATCTCCATCTTGATAAATGG
Sample 123 >A28_A1
TGGCAACGACCCCTTGTCACAGTAAAAATAGGAGGACAGTTAAAAGAAGCTCTATTGGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAAATTTGCCAGGGAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAGCAGTATGATCAGATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCWGTCAATATAATTGGAAGAAAT
ATGTTGACTCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTACCAGTAA
CATTAAAGCCAGGAATGGATGGCCCAAGAATTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGCCAAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAGTATTTGCTATAAAGAGAAAAGATGGCACTAAATGGAGGA
AATTAGTAGATTTTAGAGAGCTCAATAAAAGAACTCAGGATTTCTGGGAAGTTCAATTAGGAA
TACCGCATCCAGCGGGTTTGCAAAAGAACAAATCAGTAACAGTACTAGATGTGGGGGACGCA
TATTTTTCAGTTCCTTTAGATGAAAACTTTAGAAAGTATACTGCGTTCACCATACCTAGTATAA
ACAATGAGACACCAGGAATCAGGTATCAATACAATGTGCTTCCACAAGGATGGAAAGGATCA
CCAGCAATATTTCAGAGTAGCATGACAAAGATCTTAGAGCCCTTTAGATCACAAAATCCAGAA
ATAATTATCTGTCAATACGTGGATGACTTGTATGTAGCATCTGATTTAGAAATAGGGCAGCAT
AGAGCAAAAGTAGAGGAGTTGAGAGCTCATCTATTGAGCTGGGGATTTACTACACCAGATGA
AAAGCATCAGAAAGAACCTCCATTTCTCTGGATGGGATATGAACTCCATCCTGACAAATGG
Sample 124 >A30rd_A1
TGGCAGCGACCCCTCGTTACAGTAAAGATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTACTAGAAGAAATAAATTTACCAGGAAGATGGAAACCAAAAATG
ATAGGGGGAATTGGAGGCTTTATCAAAGTAAAACAGTATGATCAAATACCAGTAGAAATCTG
TGGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAA
ATTTGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTCCTATTGAAACTGTACCAGT
AAAATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAA
ATAAAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGACAAATTTCAAAAATTGG
GCCTGAAAATCCATACAATACTCCAATATTTGTAATAAAGAGAAAAGATGGCACTAAATGGA
GGAAATTAGTAGATTTCAGAGAACTTAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTA
GGAATACCTCATCCTGGAGGGTTAGAAAAGAGAAAATCAGTAACAGTACTAGATGTGGGTGA
TGCATATTTTTCAGTTCCATTACATGAAGACTTTAGAAAGTATACAGCATTCACCATACCTAGT
ACAAACAATGAGACACCAGGAGTTAGATATCAATACAATGTGCTTCCACAAGGATGGAAAGG
ATCACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAACCTTTTAGACAGCAAAATCC
AGAAATGGTTATCTGTCAATACATGGATGATTTGTATGTAGCATCTGATTTAGAAATAGGGCA
ACATAGAGCAAAAATAGAGGAGTTGAGAGCTCATTTATTGAAGTGGGGACTTACCACACCAG
ACAAAAAGCATCAGAAAGAACCTCCTTTCCGTTGGATGGGATATGAGCTCCATCCTGACAAGT
GG
Sample 125>A57_A1
TGGCAACGACCTCTTGTCACAGTAAAAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTTGAAGACATAAATTTGCCAGGAAAATGGAAACCAAGAATGA
TAGGGGGAATTGGAGGTTTCATTAAAGTAAAACAGTATGATCAGATACTTATAGACATTTGTG
GAAAAAAGGCTGTAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGCCCTATTGAAACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGGCCAAGGGTTAAACAATGGCCCTTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATATGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGACC
TGAAAACCCATATAATACTCCAATATTTGCAATAAAGAAAAAAGATAGCACCAAATGGAGAA
AATTAGTAGATTTCAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAGGTTCAATTAGGA
ATACCGCATCCAGGGGGTTTAAAAAAGAATAAATCAGTAACAGTACTGGATGTGGGGGACGC
ATATTTTTCAGTTCCTTTAGATGAAAGCTTTAGGAAATATACTGCGTTCACCATACCTAGTACA
AACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAATCCAG
AAATAGTTATCTATCAATACGTGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGACAGC

ATAGAACAAAAATAGAAGAGTTGAGAGCTCATCTATTGAGTTGGGGACTTACTACACCTGAC
AAAAAGCATCAAAAAGAACACCCATTCCTTTGGATGGGATATGAGCTCCTTCCTGACAAGTGG

Sample 126 >A65_A1
TGGCAACGACCCCTTGTTACAGTAAAAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGAGGAATTGGAGGTTTCATCAAAGTAAAACAGTATGATCAGATACTTATAGAAATTTGTG
GGAAAAAAGCTATAGGTACAGTATTAGTAGGACCCACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTACCAGTAA
AATTAAAACCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
CAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGGAAAATTTCAAGAATTGGGCC
CGAAAATCCATACAATACTCCAATATTTGCTATAAAGAGAAAAGACAGCACTAAATGGAGGA
AATTAGTAGATTTCAGAGAGCTCAATAAAAGAACTCAAGATTTTTGGGAAGTTCAATTAGGAA
TACCGCATCCAGCGGGGTTAAAGAAGAAGAAATCAGTAACAGTACTAGATGTGGGGGACGCA
TATTTTTCAGTTCCTTTAGATGAAAGTTTTGGGAAATATACTGCATTCACCATACCTAGTACAA
ACAATGAGACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCA
CCAGCAATATTCCAGAGTAGCATGACAAAAATTTTAGAGCCCTTCAGATCAAAAAATCCAGA
AATAATTATCTATCAATACGTGGATGACTTGTTGGTAGGGTCTGATTTAGAAATAGGGCAACA
TAGAACAAAAATAGAGGAGTTAAGAGCTCATCTGTTAAAATGGGGATTTACTACCCCAGACA
AAAAGCATCAGAAAGAACCACCATTTCTTTGGATGGGGTATGAACTCCACCCTGATAAGTGG
Sample 127>A71_A1
TGGCAACGACCTCTCGTCACAGTAAAAATAGGGGGGCAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGACACAGTATTAGAAGATATAGATTTGCCAGGAAAATGGAAACCAAAGATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGGAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGATTAAACAATGGCCATTGACAGAAGAAAAAATA
AAGGCATTAACAGAGATTTGTGCAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCGATAAAGAAAAAAGATAGCACTAAATGGAGGA
AAGTAGTAGATTTCAGAGAGCTCAATAAAAGGACACAAGATTTTTGGGAAGTTCAATTAGGA
ATACCGCACCCAGCGGGACTAAAAAAGAACAAGTCAGTAACAATTCTAGATGTGGGAGACGC
ATATTTTTCGGTCCCTTTATATGAAGACTTTAGAAAATATACTGCATTCACCATACCTAGTACA
AATAATGAGACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTTCAGAGTAGCATGACAAAGATCTTAGAACCCTTTAGATCACAAAATCCAGA
GATAGTTATCTATCAATACGTGGATGACTTATATGTAGGATCTGATTTAGAAATAGGGCAGCA
TAGAACAAAAATAGAAGAATTAAGAGCTCATCTACTRAGCTGGGGATTTACTACACCAGACA
AAAAACATCAAAAAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCACCCTGACAAGTGG
Sample 128 >A10r_A1
TGGCAACGACCCCTTGTTACAATAAAAATAGAGGGACAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGTCAGGAAAATGGAAGCCAAGAATGA
TAGGGGGAATTGGAGGTTTTATCAAGGTAAAACAATATGATCAGATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGCACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACCCAGATTGGCTGTACTCTAAATTTCCCAATTAGTCATATTGAGACTGTGCCAGTAA
AATTAAAACCAGGAATGGATGGCCCAAAGGTTAAACAGTGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTGCAGAAATGGAAAAGGAAGGAAAGATTACAAAAATTGGGCC
TGACAATCCATATAATACTCCAATATTTGCAATAAAGAGAAAAGACAACACTAAATGGAGGA
AATTAGTAGATTTCAGGGAACTCAATAAAAGAACTCAAGATTTTTGGGAAGTTCAATTGGGTA
TACCGCATCCAGCGGGCTTAAAAAAGAAGAAATCAATAACAGTACTAGATGTGGGGGACGCA
TTTTTTTCAATTCCTTTAGATGAAAGCTTTAGACAGTATACTGCGTTTACCATACCTAGTACTA
ACAATGAGACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAGGGATCA
CCAGCAATATTTCAGTATAGCATGACAAAAATCTTAGAGCCCTATAGATTAAAAAATCCAGAC
ATAGTTATCTACCAGTACGTGGATGACTTGTTAGTAGGATCTGATTTAGAAATAGGGCAGCAT

AGAGCAAAAATAGAAGAGCTAAGAGATCATTTATTGAAATGGGGATTAACTACGCCAGACAA
AAAGCATCAGAAAGAACCCCCATTTATTTGGATGGGATATGAACTCCATCCTGACAAATGG
Sample 129 >071_D
TGGCAACGACCCCTTGTCACAGTAAAAATAGGGGGACAACTAAAAGAAGCTCTGTTAGATAC
AGGAGCAGATGACACAGTATTAGAAGAAATAAATTTGCCAGGAAAATGGAAGCCAAAAATG
ATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAAATTCCTGTAGAAATCTGT
GGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAA
TTTGTTGACTCAGATTGGGTGCACTTTAAATTTTCCAATTAGTCCTATTGAAACTGTACCAGTA
AAATTAAAGCCAGGGATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAAT
AAAAGCACTAATAGAAATTTGTGCAGAACTGGAAAAGGAGGGGAAAATTTCAAGAATTGGGC
CTGAAAACCCATATAATACTCCAATATTTGCCATAAAGAAAAAGGACAGCACTAAGTGGAGA
AAATTAGTAGATTTTAGAGAACTTAATAAAAGAACTCAGGATTTCTGGGAAGTTCAACTAGGA
ATACCACACCCTGCAGGGCTAAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATGC
ATATTTTTCAGTACCCTTAGATGAAGACTTTAGAAAATACACTGCATTTACCATACCTAGTATA
AACAATGAGACACCGGGAGTTAGATATCAGTACAATGTGCTCCCACAAGGGTGGAAAGGATC
CCCGGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAACCTTTTAGAAAACAAAATCCAG
ACATAGTTATCTATCAATACGTGGATGATTTGTTAGTGGGATCTGACTTAGAAATAGGACAGC
ATAGAATAAAAGTAGAGGAATTAAGGGAACACCTATTGAAGTGGGGATTTTTCACACCAGAC
AAAAAGCATCAGAAAGAACCTCCATTTCTGTGGATGGGTTATGAACTCCATCCTGAT
Sample 130 >64r _D
TGGCAGCGACCCCTAGTCACAGTAAAGATAGGAGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAAATTTACCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGCTTTATCAAAGTAAGACAGTATGATCAAATTCTAGTAGAAATCTGTG
GACATAAAGCTATCGGTACAGTATTAGTAGGACCTACACCTGTCAACATCATTGGAAGAAATT
TGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTACCAGTAAA
ATTAAAGCCAGGGATGGATGGTCCAAAAGTTAAACAATGGCCATTGACAGCAGAAAAAATAA
AGGCACTAACAGAAATTTGTACAGAAATGGAGAAGGAAGGAAAAATTTCAAAAATTGGGCCT
GAAAATCCATATAATACTCCAATATTTGCAATAAAGAAAAAAGACAGTACTGAGTGGAGAAA
GTTAGTAGATTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAATTCAACTAGGAAT
ACCACATCCTGCAGGGTTAGAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCTT
ATTTTTCAGTTCCTTTATATGAAGACTTTAGAAAATATACTGCATTCACTATACCTAGTATAAA
TAATGAGACACCAGGAATTAGATATCAGTACAATGTACTTCCACAAGGATGGAAAGGATCAC
CAGCAATTTTCCAAAGTAGCATGACAAGAATCTTAGAACCTTTTAGAAAACAAAATCCAGAA
ATGATTATTTGTCAATACATAGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAGCAT
AGAATGAAAATAGAGGAATTAAGGGAACATCTGTTGAAGTGGGGATTTACCACACCAGACAA
AAAGCATCAGAAAGAACCTCCATTTCGTTGGATGGGTTATGAACTCCATCCTGATAAATGG
Sample 131>48_D
TGGCAACGACCCCTCGTCACAGTAAAGATAGGGGGACAGCTAAAGGAAGCTCTCTTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATGAATCTGCCAGGAAAATGGAAACCAAAAATG
ATAGGGGGAATTGGAGGCTTTATCAAAGTAAGACAATATGAACAAATACCTGTAGAAATTAG
TGGACATAAAGCTATAGGTACAGTGTTAGTAGGACCTACACCTGTCAACATAATTGGAAGAA
ATTTGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTCCTATTGACACAGTACCAGT
AAAATTAAAGCCAGGGATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAA
ATAAAAGCACTAATAGAAATTTGTACAGATATGGAAAAGGAAGGAAAAATTTCAAGAATTGG
GCCTGAAAATCCATATAATACTCCAATATTTGCCATAAAGAAAAAAGACAGTACTAGGTGGA
GAAAATTAGTAGATTTCAGGGAACTTAATAAAAGAACTCAAGACTTCTGGGAAATTCAACTA
GGAATGCCACATCCTGCAGGGCTAAAAAAGAACAAATCAGTAACAGTATTGGATGTGGGTGA
TGCATATTTCTCAGTTCCTTTATATGAGGACTTTAGAAAATATACTGCATTCACCATACCTAGT
CTGAACAATGAGACACCAGGAATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGG
ATCACCAGCAATATTTCAAAGTAGCATGACAAAAATCTTAGACCCTTTTAGAAAACAAAATCC
AGACATAGTTATTTATCAATACGTGGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCA
GCATAGAGCAAAAATAGAGGAATTAAGGGAACACCTCTTGAAATGGGGGTTTACCACACCAG

ACCAAAAGCATCAGAAAGAACATCCATTTCACTGGATGGGCTATGAACTCCATCCTGATCAAT
GG
Sample 132 >088_A1
TGGCAACGACCTATTGTTACAGTGAGAATAGGGGGACAGTTAAAAGAAGCTCTATTAGACAC
AGGAGCAGATGATACAGTATTAGAAGATATAAATTTACAAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTCATCAAAGTAAGACAGTATGATCAGGTACCTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGGCCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTYCCAATTAGCCCTATTGAGACTGTACCAGTAA
AAYTAAAACCAGGAATGGATGGACCAAGGGTTAAACAATGGCCATTGACAGAAGAAAAAAT
AAAAGCATTAACAGAAATTTGTACAGAAATGGAGAAGGAGGGAAAAATTTCAAAAATTGGGC
CTGAAAATCCATACAATACACCAATATTTGCAATAAAGAAAAAAGATGGCACTAAATGGAGA
AAGTTAGTGGATTTCAGAGAGCTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTGGGA
ATACCACATCCAGGGGGCTTAAAAAAGAACAAATCAGTAACAGTACTGGATGTGGGGGACGC
ATATTTTTCAGTTCCTTTAGATGAAGACTTTAGAAAATATACTGCGTTCACCATACCTAGTATA
AATAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTGCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAGAATCTTAGAGCCCTTTAGGAAGAGGAATCCAG
AAATAATTATCTATCAATACGTGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGC
ATAGAGCAAAAATAGAGGAATTGAGAGCTCATTTATTGAGCTGGGGACTTACTACACCAGAC
AAAAAACATCAGAAAGAACATCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAATGG
Sample 133>138_A1
TGGCAACGACCAGTAGTCACAGTAAAAGTAGGGGGACAGCTAAAAGAAGCTCTAATAGATAC
AGGAGCAGATGATACAATAATAGAAGACATAAATTTGCCAGGGAAATGGAAACCAAGAATA
GCAGGGGGACTTGGAGGTTTCATCAAAGTAAAACAGTATGATCAGATAGTCATAGAAATTTGT
GGAAAAAAGGCTATAGGTTCAGTATTAGTAGGACCCACACCAATGAACATAATTGGAAGAAA
CATATTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTACCAGTA
AAACTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGCAGAAAAAAT
AAAAGCACTAACAAAAATATGTGCAGAAATGGAAAAGGAAGGACAAATTTCAAAAATTGGG
CCTGAAAATCCATACAATACTCCAATATTTGCAATACAGAAAAAGAACAGGATAGAATGGAG
GAAATTAGTAGACTTCAGAGAGCTCAATAAGAGAACACAAGATTTTTGGGAAGTTCAATTGG
GAATACCGCATCCAGCGGGCCTAAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGAT
GCATTTTTTTCAGTTCCTTTACATGAAGATTTTAGAAAGTATACTGCGTTCACCATACCTAGTA
CAAACAATGAGACACCAGGAATTAGGTATCAGTACAATGTGCTTCCTCAGGGATGGAAAGGA
TCACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAACCCTTTAGATCAAGAAATCCA
GAAATAATCATCTGTCAATACGTGGATGACTTGTATGTAGGATCTGATTTGGAAATAGGACAG
CATAGAGCAAAAATAGAAGAGCTGAGAGCTCATCTGTTGAGCTGGGGATTTACTACACCAGA
CAAAAAGTATCAGAAGGAACCTCCATTCCAATGGATGGGATATGAGCTCCATCCTGACAAGT
GG
Sample 134 >A8a_BC
TGGCAACGACCCCTCGTCACAATAAAGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATGAGTTTGCCAGGAAGAAGGAAACCAAAAATG
ATAGGGGGAATTGGAGGTTTTATCAAAGTAAGCCAGTATGATCAGATACTCATAGAAATTTGG
GGAAATAAAGCTATAGGTACAGGATTTGTAGGCCCTACACCTGTCAACATAATTGGAAGAAA
TCTGTTGACTCAGATTGGGTGCACAATAAATTTTCCCAATAGCCCTATTGAGACTGTACCAGTA
AAATTAAAGCCAGGAATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAGAAAAT
AAAAGCATTAACAGCAATTTGTGAAGATCTGGAGAAGGAAGGAAAAATTACAAGAATTGGGC
CTGAAAATCCATATAACACTCCAATATTTGCTATAAAAAAGAAGGACAGTACCAAGTGGAGA
AAATTAGTAGATTTCAGGGAACTCAATAAAAGAACTCAAGAATTTTGGGAAGTTCAATTAGG
GATACCACACCCAGCAGGGTTAAAAAAGAGCAAATCAGTGACAGTGCTGGATGTGGGGGATG
CATATTTTTCAGTGCCTTTGGATGAAGGCTTTAGGAAATATACTGCATTCACAGTACCTAGTAT
AAATAATGAAACACCAGGCATTAGATATCAATATAATGTGCTTCCACAAGGATGGAAAGGAT
CACCAGCAATATTTCAGAGTACTATGACAAAAATCTTAAGGCCCTTTAGGGAACAAAACCCAG
ACATAGTTATCTATCAATACGTGGATGACTTGTATGTAGGATCTGACTTAGAAATAGGGCAGC

ATAGAGCAAAAATAGAGGAATTAAGACAACATCTATTGAAGTGGGGGTTTTACACACCAGAC
AAGAAGCATCAGAAAGAACATCCATTTCTGTGGATGGGGTATGAACTCCATCCTGACAAATG
G
Sample 135 >A9Br_A1
TGGCAACGACCCCTCGTCACAATAAAGATAGGGGGGCAACTAAAGGAAGCTCTAATAGATAC
AGGAGCAGATGATACAATAATAGAAGACATAAATTTGCCAGGGAAATGGAAACCAAGAATA
GCAGGGGGACTTGGAGGTTTCATCAGAGTAAAACAGTATGATCAGATAGTTATAGAAATTTGT
GGAAAAAAGGCTATAGGTTCAGTATTAGTAGGACCCACACCAGTGAACATAATTGGAAGAAA
CATATTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAGACTGTACCAGTA
AAACTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGTCAGCAGAAAAAAT
AAAAGCACTAACAGAAATATGTGCAGAAATGGAAAAGGAAGGACAAATTTCAAAAATTGGG
CCTGAAAATCCATACAATACTCCAGTATTTGCAATAAAGAAAAAGAACAGCACTAGATGGAG
AAAATTAGTAGACTTCAGAGAGCTCAATAAGAGAACACAAGATTTTTGGGAAGTTCAATTGG
GAATACCGCATCCAGCGGGCCTAAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGAT
GCATTTTTTTCAGTTCCTTTACATGAAGATTTTAGAAAGTATACTGCGTTCACCATACCTAGTA
CAAACAATGAGACACCAGGAATTAGGTATCAGTACAATGTGCTTCCTCAGGGATGGAAAGGA
TCACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAACCCTTTAGAACAAGAAATCCA
GAAATAATCATCTGTCAATACGTGGATGACTTATATGTAGGATCTGATTTGGAAATAGGACAG
CATAGAGCAAAAATAGAAGAGCTGAGAGCTCATCTGTTGAGCTGGGGGTTTTTTACACCAGAC
CAAAAATATCAGAAGGAACCTCCATTCCAATGGATGGGATATGAGCTCCATCCTGACAAGTG
G
Sample 136 >A53_D
TGGCAACGACCCCTCGTCACAATAAAGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATGAATTTACCAGGAAAATGGAAACCAAAAATGA
TAGGAGGAATTGGAGGTTTTATAAAAGTAAGACAATATGATCAAATAAGTGTAGAAATTTGT
GGAAAAAAGACTATAGGTACAGTATTAATAGGACCTACACCTGTCAACATAATTGGAAGGAA
CTTGTTGACTCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAAACTGTACCAGTA
AAATTAAAGCCAGGGATGGATGGCCCAAAAGTTAAACAATGGCCATTGTCAAAAGAAAAAAT
AAAAGCATTAACAGAAATTTGTTCAGATATGGAAAAGGAAGGAAAAATATCAAGAATTGGGC
CTGAAAATCCATACAATACTCCAATATTTGCTATAAAGAGAAAAGACAAAACTAAATGGAGA
AAATTAGTAGATTTCAGAGAACTTAATAAAAGAACTCAAGACTTTTGGGAAATTCAATTAGGA
ATACCACATCCTGCAGGGTTGGAAAAGAAAAAATCAGTAACAATACTGGATGTGGGGGATGC
ATATTTTTCAGTTCCCTTGGATGAAAATTTTAGAAAATATACAGCATTTACCATACCTAGTATA
AACAATGAGACACCAGGAATTAGATATCAGTACAATGTACTTCCACAAGGATGGAAAGGATC
ACCAGCAATATTCCAAGATAGCATGACGAAAATTTTAAGACCCTTTAGAGAACAAAATCCAG
ACATAATTATCTGTCAATACGTGGACGATTTGTATGTAGCGTCTGACTTAGAAATAGGGCAGC
ATAGAATAAAAATAGAAGAATTAAGGAAACACCTGTTGAAGTGGGGATTTACCACACCAGAC
ACAAAGTATCAGAAAGAGCCCCCATTTCTTTGGATGGGTTATGAACTTCATCCTGATAAATGG
Sample 137 >A57B_A1
TGGCAACGACCTCTTGTCACAGTAAAAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTTGAAGACATAAATTTACCAGGAAAATGGAAACCAAGAATGA
TAGGGGGAATTGGAGGTTTCATTAAAGTAAAACAGTATGATCAGATACTTATAGACATTTGTG
GAAAAAAGGCTGTAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAAACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGGCCAAGGGTTAAACAATGGCCCTTGACAGAAGAAAAGATA
AAAGCATTAACAGAAATATGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAACCCATATAATACTCCAATATTTGCAATAAAGAAAAAAGATAGCACCAAATGGAGAA
AATTAGTAGATTTCAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAGGTTCAATTAGGA
ATACCGCATCCAGGGGGTTTAAAAAAGAATAAATCAGTAACAGTACCGGATGTGGGGGACGC
ATATCTTTCAGTTCCCATTGATGAAAACTAAAGGAAATAACTGCGTTCACCATACCTAGTACA
AACAATGAGACACCAGGAACCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAATCCAG

AAATAGTTATCTATCAATACGTGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGACAGC
ATAGAACAAAAATAGAAGAGTTGAGAGCTCATCTATTGAGTTGGGGACTTACTACACCTGAC
AGAAAGCATCAAAAAGAACACCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTG
G
Sample 138 >A72 _D
TGGCAACGACCCCTCGTCACAATAAAGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATGAATTTGTCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAATATGAACAAATACCTGTAGAAATTTGTG
GACATAAAGCTATAGGTACAGTGTTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATC
TGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTCCTATTGACACAGTACCAGTAAA
ATTAAAGCCAGGGATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAATAA
AAGCACTAATAGAAATTTGTACAGATATGGAAAAGGAAGGAAAAATTTCAAGAATTGGGCCT
GAAAATCCATATAATACTCCAATATTTGCCATAAAGAAAAAAGACAGTACTAGGTGGAGAAA
ATTAGTAGATTTCAGGGAACTTAATAAAAGAACTCAAGACTTCTGGGAAGTTCAACTAGGAAT
ACCACATCCCGCAGGGCTAAAAAAGAACAAATCAGTAACAGTATTGGATGTGGGTGACGCAT
ATTTCTCAGTTCCTTTATATGAGGACTTTAGAAAATATACTGCATTCACCATACCTAGTTTGAA
CAATGAGACACCAGGAATTAGATATCAGTACAATGTGCTTCCACAAGGATGGAAAGGATCAC
CAGCAATATTTCAAAGTAGCATGACAAAAATCTTAGACCCTTTTAGAAAACAAAATCCAGACA
TAGTTATTTATCAATACGTGGATGATTTGTATGTAGGATCTGACTTGGAAATAGGGCAACATA
GAGCAAAAATAGAGGAATTAAGGGAACACCTCTTGAAATGGGGGTTTACCACACCAGACCAA
AAGCATCAAAAAGAACATCCATTTCATTGGATGGGCTATGAACTCCATCCTGATCAATGG
Sample 139>A42r_D
TGGCAACGACCCCTTGTCACAGTAAAGATAGGGGGACAGCTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAAATATGTGTAGAAATCTGTG
GATATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATT
TATTGACTCAGATTGGTTGCACTCTAAATTTTCCCATTAGTCCTATTGAAACTGTACCAGTAAA
ATTAAAACCAGGGATGGATGGCCCAAAAGTTAAACAGTGGCCGTTGACAGAAGAAAAAATAA
AAGCACTAGTAGAAATTTGTACAGAAATGGAAAAGGAAGGGAAAATTTCAAGAATTGGGCCT
GAAAATCCATACAATACTCCAATATTTGCCATAAAGAAAAAAGACAGTACTAAGTGGAGAAA
ATTAGTAGATTTCAGAGAACTGAATAAGAGAACGCAAGACTTCTGGGAAGTTCAACTAGGAA
TACCACATCCTGCAGGGCTAAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCA
TATTTTTCAGTTCCCCTAGATGCAAACTTCAGAAAATATACTGCATTCACCATACCTAGTACAA
ACAATGAGACACCAGGAATTAGATATCAGTACAATGTGCTTCCACAAGGATGGAAAGGATCA
CCGGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAACCCTTTAGAAAACAAAATCCAGA
AATAGTTATCTATCAATACGTGGATGATTTGTATATAGGATCTGACTTAGAATTAGGGCAGCA
TAGAACGAAAATAGAGGAATTAAGGGAACACCTATTGAGGTGGGGATTTACTACACCAGACA
AAAAGCATCAGAAAGAACCTCCGTTTCTTTGGATGGGTTATGAACTCCATCCTGACAAATGG
Sample 140 >002_A1B
TGGCAACGACCCCTCGTCACAATAAAGGTAGGGGGACAGCAAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAGATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAAATACCTATAGAAATTTGTG
GAAGAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGCTTGGGTGTACTTTAAATTTCCCAATTAGTCCTATTGAAACTGTACCAGTAA
AATTAAAACCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTACAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCAATAAGGAAAAAAGATAGTACTAAATGGAGGA
AGCTAGTAGATTTTAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAGGTTCAATTGGGA
ATACCGCATCCAGCGGGCCTAAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGACGC
ATATTTTTCAGTTCCTTTAGATGAAAGCTTTAGAAAGTATACTGCGTTCACCATACCTAGTACA
AACAATGCAACACCAGGAATCAGATATCAGTACAATGTACTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTTCAGAGTAGCATGACAAAAATCTTAGATCCCTTTAGATCAAAAAATCCAGA

AATAATTATCTATCAATATATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGCA
TAGAGCAAAAGTAGAAGAGTTGAGAGCTCATCTATTGAGATGGGGATTTACTACACCAGATA
AAAAGCATCAGAAAGAACCCCCATTCCTTTGGATGGGTTATGAGCTCCATCCTGACAAGTGG
Sample 141>005_A1B
TGGCAGCGACCCCTTGTCACAGTAAGAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAGAATGA
TAGGGGGAATCGGAGGGTTCATCAAGGTAAAACAGTATGATCAGATACTTATAGAAATTTGT
GGAAAAAAGGCTATAGGTACCGTATTGGTAGGACCTACACCTGTCAACATAATTGGAAGGAA
CATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTAATACTGTACCCGTA
AAATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAAT
AAAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGC
CTGAAAATCCATACAATACTCCAATTTTTGCTATAAAGAAAAAGGACAGCACTAAATGGAGA
AAATTAGTAGATTTCAGAGAACTCAATAAAAGAACTCAAGACTTTTGGGAAAGTTCAATTAGG
AATACCGCATCCCGCGGGCTTAAAAAAGAAAAAATCAGTCACAGTACTAGATGTGGGGGATG
CATACTTTTCAGTTCCCTTAGATGAAGGCTTTAGAAAGTATACTGCATTCACCATACCTAGTAC
AAACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGAT
CACCGGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATCGAAAAATCCAG
AGATAATTATATATCAATACATGGATGACTTGTATGTAGGATCCGATTTAGAAATAGGGCAGC
ATCGAACAAAAATAGAGGAGCTGAGACAACATCTGTTGAGGTGGGGACTTACCACACCAGAC
AAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATTAACTCCATCCTAACAAATGG
Sample 142 >A30 _A1D
TGGCAACGACCCCTTGTCACAGTAAGAGTAGGGGGACAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAGATTTGCCAGGAAGATGGAAACCAAGAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAGGTACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTGTTAGTAGGGCCCACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGATTGGCTGTACTTTAAATTTCCCAATTAGTCCTATTGAAACTGTACCAGTAA
AATAAAAACCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAAT
AAAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAGATTGGGC
CTGAAAATCCATACAATACTCCAATATTTGCAATAAAGAAAAAAGACAGCACTAAATGGAGG
AAATTAGTAGATTTCAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGG
GATACCGCACCCAGCGGGCTTAAAAAAGAAGAAATCAGTAACAGTACTGGATGTGGGGGACG
CATATTTTTCAGTTCCTTTAGATGAAAGCTTTAGAAAGTATACTGCGTTCACCATACCCAGTAT
AAACAATGAGACACCAGGAATAAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGGT
CACCGGCAATATTCCAGAGCAGTATGACAAAAATCTTAGAGCCCTTTAGGTTACAAAATCCAG
ACATGACTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGC
ATAGAGCAAAAATAGAAGAATTAAGAGCTCATCTATTGAGCTGGGGATTTACTACACCAGAC
AAAAAGCATCAGAAAGAACCCCCATTCCTTTGGATGGGATATGAGCTTCATCCTGACAAGTGG
Sample 143 >A99M_A1
TGGCAGCGACCCCTTGTTACAGTAAGAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAGATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGGTTCATCAAGGTAAAACAATATGATCAGATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTGGGGAGACCTACACCTGTTCACATAATTGGAAGAAAT
CTGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAAACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCGAGGGTTAAACAATGGCCATTATCAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTACAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAGTATTTGCAATAAAGAAAAAGGATAGCACTAAATGGAGGA
AACTAGTAGATTTCAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGGA
ATCCCGCATCCAGCGGGCTTAAAAAAGAAGAAATCAGTAACAGTGCTAGATGTGGGGGACGC
ATATTTCTCAGTTCCTTTAGATGAAGGTTTTAGAAAATATACTGCATTCACCATACCTAGTACA
AACAATGAAACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATTTTAGAACCCTTTAGAGCAAAGAATCCAG
AAATAATTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGC

ATAGAACAAAAATAGAGGAGTTAAGAGATCATCTATTGAAGTGGGGATTTACTACACCAGAC
AAAAAGCATCAGAAAGAACCCCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 144 >072r_A1
TGGCAACGACCTGTAGTCACAGTAAGAATAGAGGGACAACTAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGAAATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTGGTAGGACCTACACCTGTCAACATAATTGGAAGGAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCCATTGATACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGCACAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGACGGCACTAAATGGAGAA
AATTAGTAGATTTTAGAGAGCTTAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGGA
TACCGCATCCAGCGGGTTTGAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGACGCC
TATTTTTCAGTTCCCTTAGATGAAGACTTTAGAAAATATACAGCATTCACCATACCTAGTACAA
ACAATGAGACACCAGGAATAAGGTATCAGTACAATGTACTTCCACAGGGATGGAAAGGATCA
CCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCTTTTAGAGCAAAAAATCCAGA
AATGGTAATCTATCAGTACATGGATGACTTGTATGTAGGCTCTGATTTAGAAATAGGGCAGCA
TAGAACAAAAGTAGAAGAATTAAGAAAACATCTATTGAGCTGGGGATTTACTACACCAGACA
AAAAGCATCAGAAAGAACCTCCATTTCTCTGGATGGGGTATGAACTCCATCCTGATAAGTGG
Sample 145 >A53r_A1
TGGCAGCGACCCCTTGTCACAGTGAAAATAGGGGGACAGCTAAAAGAAGCTCTATTAGATAC
AGGAGCCGATGATACAGTATTAGAAGACATAGATTTGCCAGGAAAATGGAAGCCAAAAATGA
TAGGGGGAATTGGAGGTTTCATCAAGGTAAAACAGTATGATCAGATACTTATGGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAAACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAAATTAAACAATGGCCATTGACAGAGGAAAAAATA
AAAGCATTAACAGAGATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATCGGGCC
TGAAAATCCATACAATACTCCCATATTTGCAATAAAGAAAAAAGATAGCACTAAATGGAGAA
AATTGGTAGATTTCAGAGAGCTTAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGGA
ATACCGCATCCAGCGGGCCTAAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGACGC
ATATTTTTCAGTTCCTTTAGATGAAAGCTTTAGAAAGTATACTGCATTCACCATACCTAGTATA
AACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGATCCTTTTAGAGCAAAGAATCCAGA
AATAGTTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGAGCAGCA
TAGAGCAAAAATAGAAGAGTTAAGAGCTCATCTATTGAGCTGGGGATTTACTACCCCAGACA
AGAAGCATCAGAAAGAACCCCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGG
Sample 146 >A31r_A1
TGGCAACGACCTCTTGTCACAGTAAAAATAGGGGGACAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGAGCAAGTACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTGGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT
ATGTTGACTCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGATACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTAACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGCATAGAGATGGAAAAGGAAGGGAAAATTACAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAGGATAGCACTAAATGGAGAA
AATTAGTAGATTTCAGAGAACTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTGGGAA
TACCGCATCCAGCGGGCTTAAAAAAGAAGAAATCAGTAACAGTACTAGATGTGGGGGATGCA
TATTTTTCAGTTCCCTTAGATGAAAACTTTAGAAAGTATACTGCATTTACAATACCTAGTACAA
ACAATGAGACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCA
CCAGCAATATTCCAGCATAGCATGACAAAAATCTTAGAGCCCTTTAGATAAAAAAATCCAGA
AATAATTATCTACCAATACATGGATGACTTGTATGTAGGATCTGATTTGGAAATAGGACAACA

TAGAACAAAAATAGAAGAGTTGAGAGCTCATCTATTGAGCTGGGGATTTACTACACCAGACA
AAAAGCATCAGAAGGAACCTCCATTTCTTTGGATGGGATATGAACTTCATCCAGACAAGTGG
Sample 147 >A16r_A1
TGGCAACGACCCCTCGTCACAGTAAAAGTAGGAGGACAGATAAAAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGATATAGATTTGCCAGGAAGATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAGATAACTATGGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGGAAT
ATGTTGACTCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCATATTGAGACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTAAAGAGATGGAAGAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAGTATTTGCTATAAAGAAAAAAGATAGCACTAAATGGAGAA
AGCTAGTAGATTTTAGAGAGCTCAATAAAAGAACTCAGGATTTCTGGGAAGTTCAATTAGGAA
TACCACACCCTGCAGGGTTAAAAAAGAAGAAATCAGTAACAGTACTAGATGTGGGAGACGCA
TATTTTTCAGTTCCTTTACATGAAAGCTTTAGAAAATATACTGCGTTCACCATACCTAGTATAA
ACAATGAAACACCAGGAATCAGATACCAGTACAATGTGCTTCCACAAGGATGGAAAGGGTCA
CCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAATCCAGA
GATAATTATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGCA
TAGAGCAAAAATAGAAGAGTTAAGAGCTCACCTGTTGAGCTGGGGGTTTACTACACCAGACA
AAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAGCTCCATCCTGACAAATGG
Sample 148 >050r_A1
TGGCAACGACCTCTTGTCCCAGTGAAAATAGAGGGACAGCTAAGAGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAAATTTGTCAGGAAAATGGAAACCAAGAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACTCAAATTGGTTGTACTTTAAATTTCCCAATTAGTCCTGTTGAGACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGTCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCAATAAAGAAAAAAGATAGCACTAAATGGAGAA
AACTAGTAGATTTCAGAGAACTCAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGGA
ATCCCGCATCCAGCGGGCTTAAAAAAGAAAAAATCAGTAACAGTATTAGATGTGGGGGACGC
ATATTTTTCAGTTCCTTTATATGAAGATTTTAGAAAATATACTGCGTTCACCATACCTAGTATG
AACAATGAGACACCAGGGATCAGGTATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTCCAGCATAGTATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAATCCAGA
AATAATTATTTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGAGCAGCA
TAGAGCAAAAATAGAAGAGTTGAGAGCTCATTTATTGCACTGGGGGTTTACTACACCAGACA
AAAAGCATCAGAAAGAACCTCCATTCCTTTGGATGGGATATGAACTCCATCCTGACAAGTGG
Sample 149 >A17r_A1
TGGCAACGACCCCTTGTCACAGTAAAAATAGGGGGACAGCTAAAGGAAGCTCTATTAGACAC
AGGAGCAGATGATACAGTATTTGAAGACATAAATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTCATCAAGGTAAAACAGTATGATCAGATACTTATAGAAATTTGTG
GAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCAGTCAACATAATTGGAAGAAAT
ATGTTGACCCAGATTGGTTGTACTTTAAATTTCCCAATTAGTCCTATTGAAACTGTACCAGTAA
AATTAAAGCCAGGAATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTATGGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCTATAAAGAAAAAAGACAGCACTAAATGGAGGA
AACTAGTAGATTTTAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAAGTTCAATTAGGA
ATACCGCATCCAGCGGGCCTAAAAAAGAAGAAATCAGTAACAGTACTAGATGTGGGGGACGC
ATATTTTTCAGTTCCTTTAGATGAAAGCTTTAGAAAATATACTGCATTCACCATACCTAGTACA
AACAATGAGACACCAGGAATCAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGGTC
ACCGGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAATCCAG
AAATAATTATTTATCAATACATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGC

ATAGAGCAAAAATAGAAGAATTGAGAGCTCATCTATTGAGCTGGGGGTTTACTACACCAGAT
AAAAAACATCAGAAAGAACCTCCATTTCTTTGGATGGGATATGAACTCCATCCTGATAAATGG
Sample 150 >2M _A1
TGGCAACGACCCCTCGTCACAATAAAGGTAGGGGGACAGCAAAAGGAAGCTCTATTAGATAC
AGGAGCAGATGATACAGTATTAGAAGACATAGATTTGCCAGGAAAATGGAAACCAAAAATGA
TAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAAATACCTATAGAAATTTGTG
GAAGAAAGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAC
ATGTTGACCCAGCTTGGGTGTACTTTAAATTTCCCAATTAGTCCTATTGAAACTGTACCAGTAA
AATTAAAACCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATA
AAAGCATTAACAGAAATTTGTACAGAGATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCC
TGAAAATCCATACAATACTCCAATATTTGCAATAAGGAAAAAAGATAGTACTAAATGGAGGA
AGCTAGTAGATTTTAGAGAGCTCAATAAAAGAACACAAGACTTTTGGGAGGTTCAATTGGGA
ATACCGCATCCAGCGGGCCTAAAAAAGAAAAAATCAGTAACAGTACTAGATGTGGGGGACGC
ATATTTTTCAGTTCCTTTAGATGAAAGCTTTAGAAAGTATACTGCGTTCACCATACCTAGTACA
AACAATGCAACACCAGGAATCAGATATCAGTACAATGTACTTCCACAGGGATGGAAAGGATC
ACCAGCAATATTTCAGAGTAGCATGACAAAAATCTTAGATCCCTTTAGATCAAAAAATCCAGA
AATAATTATCTATCAATATATGGATGACTTGTATGTAGGATCTGATTTAGAAATAGGGCAGCA
TAGAGCAAAAGTAGAAGAGTTGAGAGCTCATCTATTGAGATGGGGATTTACTACACCAGATA
AAAAGCATCAGAAAGAACCCCCATTCCTTTGGATGGGTTATGAGCTCCATCCTGACAAGTGG

Appendix 7:- Report for Mutations generated by Recall Software

HIV Resistance Genotype report

For research use only

Patient Details	Test Details	Physician Details
Name:	Sample ID: 8	
	Secondary ID:	
Patient ID:	Sample Date:	
Birthdate:	Report Date:	

NRTI/NtRTI Drugs	
Abacavir	Intermediate resistance
Zidovudine	Susceptible
Stavudine	Low-level resistance
Didanosine	Low-level resistance
Emtricitabine	High-level resistance
Lamivudine	High-level resistance
Tenofovir	Low-level resistance
NNRTI Drugs	
Doravirine	Low-level resistance
Efavirenz	Intermediate resistance
Etravirine	Intermediate resistance
Nevirapine	High-level resistance
Rilpivirine	High-level resistance
Relevant NRTI/NtRTI/NNRTI Mutations:	
K70E, M184I, K101E, Y181C	

PI Drugs	
Atazanavir/r	Susceptible
Darunavir/r	Susceptible
Fosamprenavir/r	Susceptible
Indinavir/r	Susceptible
Lopinavir/r	Susceptible
Nelfinavir	Susceptible
Saquinavir/r	Susceptible
Tipranavir/r	Susceptible
Relevant PI Mutations:	
None	

Appendix 8: - Report on HIV drug resistance by Stanford database

Stanford University
HIV DRUG RESISTANCE DATABASE
A curated public database to represent, store and analyze HIV drug resistance data.

HIVdb Program Report

Genotypic Resistance Interpretation Algorithm

Sierra version 3.0.5 (last updated on 2020-06-24)
HIVdb version 8.9-1 (last updated on 2019-10-25)

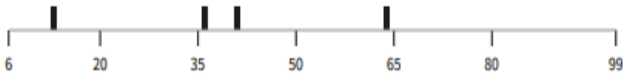
Sequence 8

Sequence summary

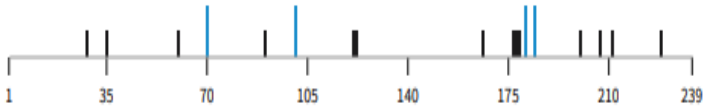
Sequence includes PR:	codons 6 - 99
Sequence includes RT:	codons 1 - 239
Subtype:	⊟ D (5.01%)
	• AF457090: **Kenya (2001); D (5.01%); best match**
	• AY322189: Kenya (1997); D (5.31%)
	• AF133821: Kenya (1993); D (5.41%)
	• AF484480: Uganda (1999); D (5.51%)
	• AY444799: United States (1999); D (5.51%)
	• AF289550: Tanzania, United Republic of (1996); C + D (CRF10_CD, 5.81%)
	• KT022409: Kenya (2005); D (5.91%)
	• AY445524: Kenya (1997); D (6.01%)
	• DQ912823: Denmark (1998); D (6.01%)
	• KT022385: Kenya (2003); D (6.01%)
PR SDRMs:	None
RT SDRMs:	K70E, K101E, Y181C, M184I

Sequence quality assessment

PR



RT



Drug resistance interpretation: PR — HIVDB 8.9-1 (2019-10-25)

PI Major Resistance Mutations:	None
PI Accessory Resistance Mutations:	None
Other Mutations:	I13V, M36I, R41K, I64V

Protease Inhibitors	
atazanavir/r (ATV/r)	Susceptible
darunavir/r (DRV/r)	Susceptible
lopinavir/r (LPV/r)	Susceptible

Mutation scoring: PR — HIVDB 8.9-1 (2019-10-25)

PI	ATV/r	DRV/r	LPV/r
Total	0	0	0

Drug resistance interpretation: RT — HIVDB 8.9-1 (2019-10-25)

NRTI Resistance Mutations:	**K70E**, **M184I**
NNRTI Resistance Mutations:	**K101E**, **Y181C**
Other Mutations:	E28A, V35T, V60I, V90I, D121Y, K122E, K166R, D177E, I178M, V179I, T200M, Q207E, R211K, L228R

Nucleoside Reverse Transcriptase Inhibitors		Non-nucleoside Reverse Transcriptase Inhibitors	
abacavir (ABC)	Intermediate Resistance	**doravirine (DOR)**	Low-Level Resistance
zidovudine (AZT)	Susceptible	**efavirenz (EFV)**	Intermediate Resistance
emtricitabine (FTC)	High-Level Resistance	**etravirine (ETR)**	Intermediate Resistance
lamivudine (3TC)	High-Level Resistance	**nevirapine (NVP)**	High-Level Resistance
tenofovir (TDF)	Low-Level Resistance	**rilpivirine (RPV)**	High-Level Resistance

RT comments

NRTI

- **K70E/G** cause low-level resistance to TDF, ABC, DDI and possibly 3TC and FTC. **K70E/G** increase susceptibility to AZT.
- **M184V/I** cause high-level in vitro resistance to 3TC and FTC and low-level resistance to ddI and ABC. However, **M184V/I** are not contraindications to continued treatment with 3TC or FTC because they increase susceptibility to AZT, TDF and d4T and are associated with clinically significant reductions in HIV-1 replication.

NNRTI

- **K101E** is a non-polymorphic primarily accessory mutation that causes intermediate resistance to NVP and RPV, low-level resistance to EFV, and potentially low-level resistance to ETR. It has a weight of 1.0 in the Tibotec ETR genotypic susceptibility score. It is associated with low-level reductions in DOR susceptibility.
- **Y181C** is a non-polymorphic mutation selected in patients receiving NVP, ETR and RPV. It reduces susceptibility to NVP, ETR, RPV, and EFV by >50-fold, 5-fold, 3-fold, and 2-fold, respectively. Although **Y181C** itself reduces EFV susceptibility by only 2-fold, it has been associated with a reduced response to an EFV-containing regimen in NNRTI-experienced patients. **Y181C** has a weight of 2.5 in the Tibotec ETR genotypic susceptibility score. Alone, it does not appear to reduce DOR susceptibility.

Other

- **V90I** is a polymorphic accessory mutation weakly selected by each of the NNRTIs. It has a weight of 1.0 in the Tibotec ETR genotypic susceptibility score but is associated with minimal, if any, detectable reduction in NNRTI susceptibility.
- **V179I** is a polymorphic mutation that is frequently selected in patients receiving ETR and RPV. But it has little, if any, direct effect on NNRTI susceptibility.

Mutation scoring: RT — HIVDB 8.9-1 (2019-10-25)

NRTI	ABC	AZT	FTC	3TC	TDF
K70E	15	-10	10	10	15
M184I	15	-10	60	60	-10
K70E + M184I	0	0	0	0	10
Total	30	-20	70	70	15

NNRTI	DOR	EFV	ETR	NVP	RPV
K101E	15	15	15	30	45
Y181C	10	30	30	60	45
K101E + Y181C	0	5	5	5	0
K101E + M184I	0	0	0	0	15
Total	25	50	50	95	105

Appendix 9: - HIV-1 Subtypes output generated by jpHMM

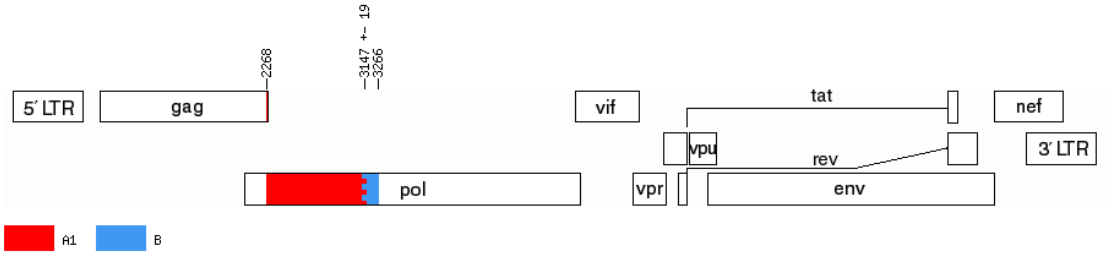
jpHMM result:

Sequence #1: >002_A1B

This sequence is related to subtype(s): **A1 B**

Fragment Start Position	Uncertainty Region Start - End	Breakpoint Interval Start - End	Fragment End Position	Fragment Subtype
Position in the original sequence [pred_recombination], [recombination_incl_UR_and_BPI], [UR_and_BPI]				
1	-	861 - 899	**894**	**A1**
895	-	-	**999**	**B**
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2268	-	3128 - 3166	**3161**	**A1**
3162	-	-	**3266**	**B**

Genome map (based on HXB2 numbering)



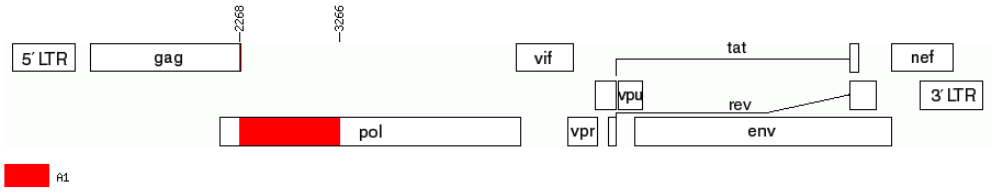
jpHMM result:

Sequence #1: >066

This sequence is related to subtype(s): **A1**

Fragment Start Position	Uncertainty Region Start - End	Breakpoint Interval Start - End	Fragment End Position	Fragment Subtype
Position in the original sequence [pred_recombination], [recombination_incl_UR_and_BPI], [UR_and_BPI]				
1	-	-	**999**	**A1**
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2268	-	-	**3266**	**A1**

Genome map (based on HXB2 numbering)



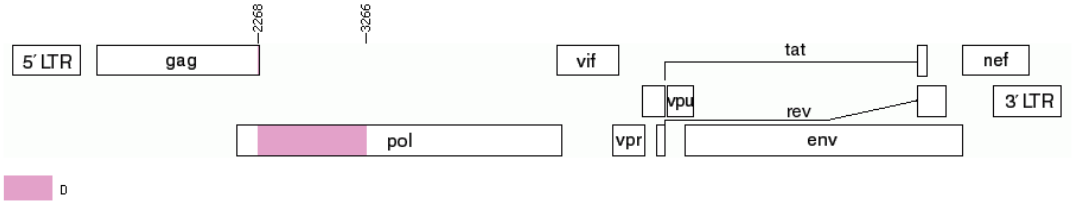
jpHMM result:

Sequence #1: >8B

This sequence is related to subtype(s): **D**

Fragment Start Position	Uncertainty Region Start - End	Breakpoint Interval Start - End	Fragment End Position	Fragment Subtype
Position in the original sequence [pred_recombination], [recombination_incl_UR_and_BPI], [UR_and_BPI]				
1	-	-	**999**	**D**
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2268	-	-	**3266**	**D**

Genome map (based on HXB2 numbering)



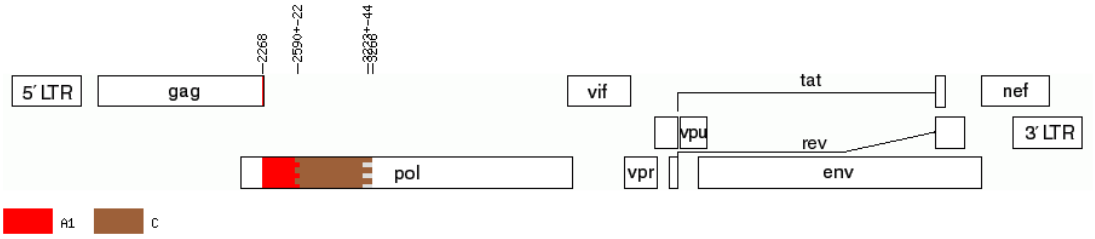
jpHMM result:

Sequence #1: >A7

This sequence is related to subtype(s): **A1 C**

Fragment Start Position	Uncertainty Region Start - End	Breakpoint Interval Start - End	Fragment End Position	Fragment Subtype
Position in the original sequence [pred_recombination], [recombination_incl_UR_and_BPI], [UR_and_BPI]				
1	-	301 - 344	**342**	**A1**
343	912 - 999	-	**999**	**C**
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2268	-	2568 - 2611	**2609**	**A1**
2610	3179 - 3266	-	**3266**	**C**

Genome map (based on HXB2 numbering)



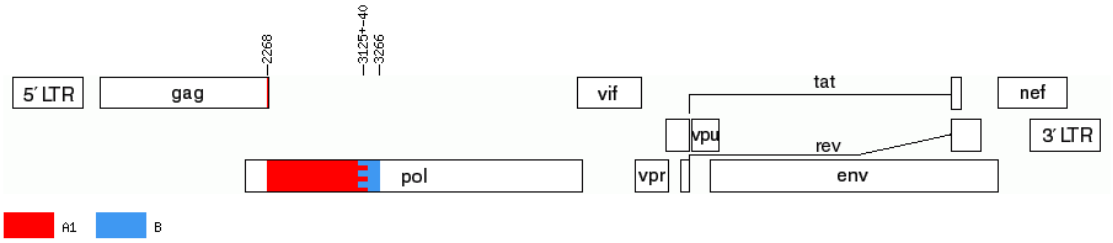
jpHMM result:

Sequence #1: >005

This sequence is related to subtype(s): **A1 B**

Fragment Start Position	Uncertainty Region Start - End	Breakpoint Interval Start - End	Fragment End Position	Fragment Subtype
Position in the original sequence [pred_recombination], [recombination_incl_UR_and_BPI], [UR_and_BPI]				
1	-	819 - 898	**895**	**A1**
896	-	-	**1000**	**B**
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2268	-	3085 - 3164	**3161**	**A1**
3162	-	-	**3266**	**B**

Genome map (based on HXB2 numbering)



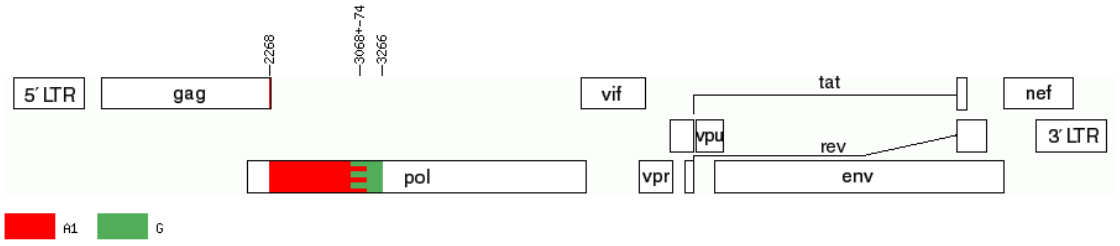
jpHMM result:

Sequence #1: >Seq12

This sequence is related to subtype(s): **A1 G**

Fragment Start Position	Uncertainty Region Start - End	Breakpoint Interval Start - End	Fragment End Position	Fragment Subtype
Position in the original sequence [pred_recombination], [recombination_incl_UR_and_BPI], [UR_and_BPI]				
1	-	727 - 874	**746**	**A1**
747	-	-	**999**	**G**
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2268	-	2994 - 3141	**3013**	**A1**
3014	-	-	**3266**	**G**

Genome map (based on HXB2 numbering)



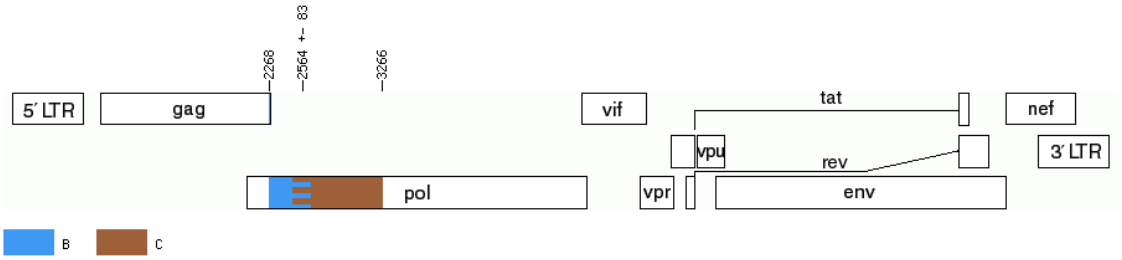
jpHMM result:

Sequence #1: **>A8**

This sequence is related to subtype(s): **B C**

Fragment Start Position	Uncertainty Region Start - End	Breakpoint Interval Start - End	Fragment End Position	Fragment Subtype
Position in the original sequence [pred_recombination], [recombination_incl_UR_and_BPI], [UR_and_BPI]				
1	-	214 - 380	**264**	**B**
265	-	-	**999**	**C**
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2268	-	2481 - 2647	**2531**	**B**
2532	-	-	**3266**	**C**

Genome map (based on HXB2 numbering)



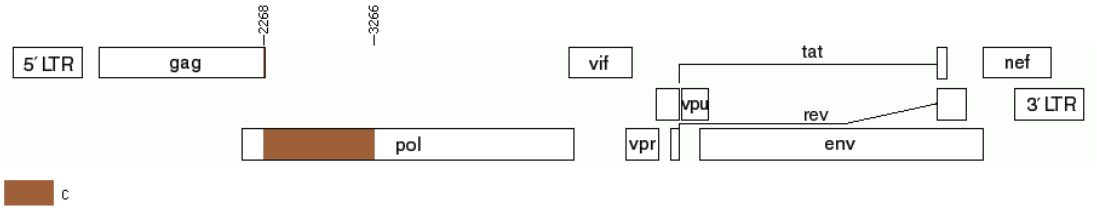
jpHMM result:

Sequence #1: **>Seq18**

This sequence is related to subtype(s): **C**

Fragment Start Position	Uncertainty Region Start - End	Breakpoint Interval Start - End	Fragment End Position	Fragment Subtype
Position in the original sequence [pred_recombination], [recombination_incl_UR_and_BPI], [UR_and_BPI]				
1	-	-	**999**	**C**
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2268	-	-	**3266**	**C**

Genome map (based on HXB2 numbering)



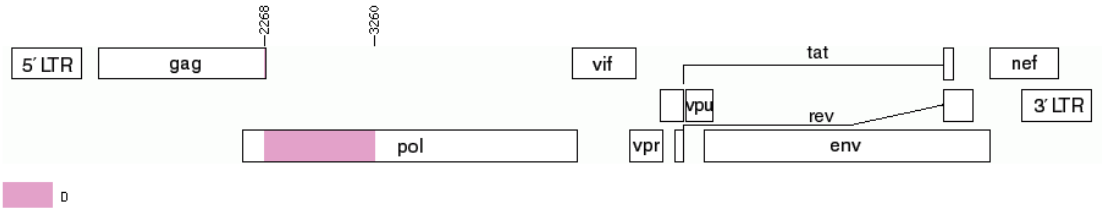
jpHMM result:

Sequence #1: >071_D

This sequence is related to subtype(s): **D**

Fragment Start Position	Uncertainty Region Start - End	Breakpoint Interval Start - End	Fragment End Position	Fragment Subtype
Position in the original sequence [pred_recombination], [recombination_incl_UR_and_BPI], [UR_and_BPI]				
1	-	-	**993**	**D**
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2268	-	-	**3260**	**D**

Genome map (based on HXB2 numbering)



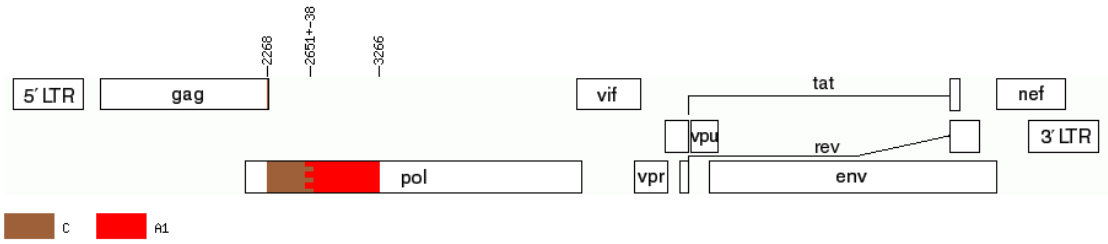
jpHMM result:

Sequence #1: >35_A1C

This sequence is related to subtype(s): **A1 C**

Fragment Start Position	Uncertainty Region Start - End	Breakpoint Interval Start - End	Fragment End Position	Fragment Subtype
Position in the original sequence [pred_recombination], [recombination_incl_UR_and_BPI], [UR_and_BPI]				
1	-	346 - 421	**397**	**C**
398	-	-	**996**	**A1**
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2268	-	2613 - 2688	**2664**	**C**
2665	-	-	**3266**	**A1**

Genome map (based on HXB2 numbering)



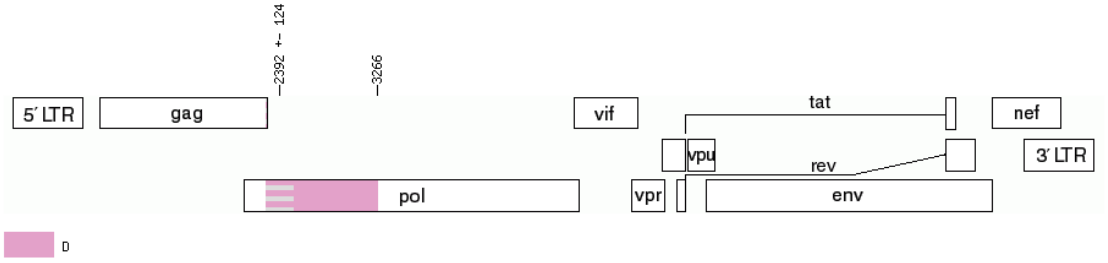
jpHMM result:

Sequence #1: >A53_D

This sequence is related to subtype(s): **D**

Fragment Start Position	Uncertainty Region Start - End	Breakpoint Interval Start - End	Fragment End Position	Fragment Subtype
Position in the original sequence [pred_recombination], [recombination_incl_UR_and_BPI], [UR_and_BPI]				
1	1 - 249	-	**999**	**D**
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2268	2268 - 2516	-	**3266**	**D**

Genome map (based on HXB2 numbering)



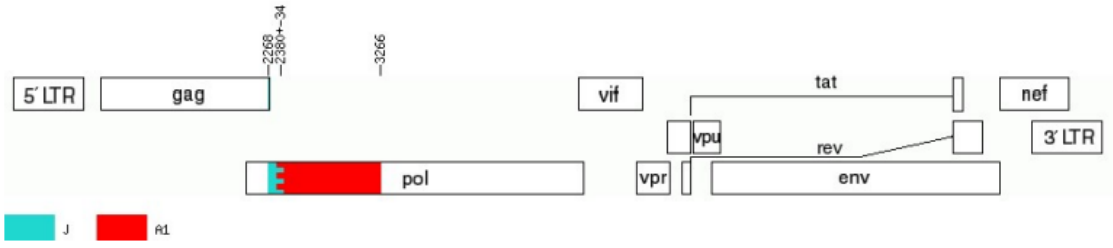
jpHMM result:

Sequence #1: >049

This sequence is related to subtype(s): **A1 J**

Fragment Start Position	Uncertainty Region Start - End	Breakpoint Interval Start - End	Fragment End Position	Fragment Subtype
Position in the original sequence [pred_recombination], [recombination_incl_UR_and_BPI], [UR_and_BPI]				
1	-	79 - 146	**94**	**J**
95	-	-	**999**	**A1**
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2268	-	2346 - 2413	**2361**	**J**
2362	-	-	**3266**	**A1**

Genome map (based on HXB2 numbering)



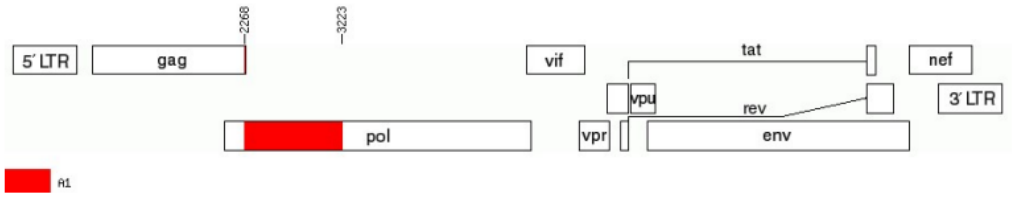
jpHMM result:

Sequence #1: >010

This sequence is related to subtype(s): **A1**

Fragment Start Position	Uncertainty Region Start - End	Breakpoint Interval Start - End	Fragment End Position	Fragment Subtype
Position in the original sequence [pred_recombination], [recombination_incl_UR_and_BPI], [UR_and_BPI]				
1	-	-	**962**	**A1**
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2268	-	-	**3223**	**A1**

Genome map (based on HXB2 numbering)



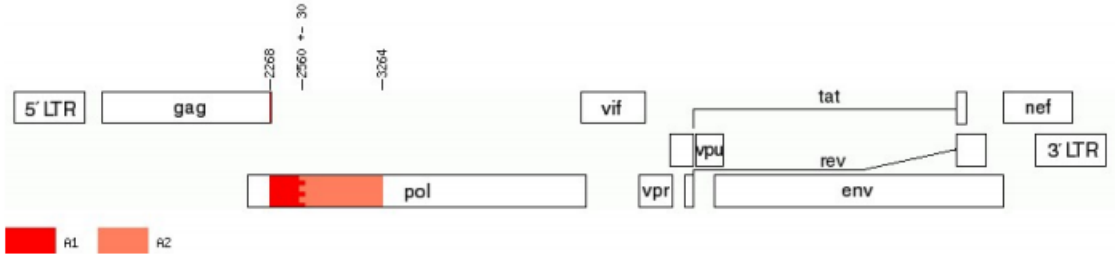
jpHMM result:

Sequence #1: >034

This sequence is related to subtype(s): **A1 A2**

Fragment Start Position	Uncertainty Region Start - End	Breakpoint Interval Start - End	Fragment End Position	Fragment Subtype
Position in the original sequence [pred_recombination], [recombination_incl_UR_and_BPI], [UR_and_BPI]				
1	-	263 - 328	**278**	**A1**
279	-	-	**999**	**A2**
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2268	-	2530 - 2590	**2545**	**A1**
2546	-	-	**3264**	**A2**

Genome map (based on HXB2 numbering)



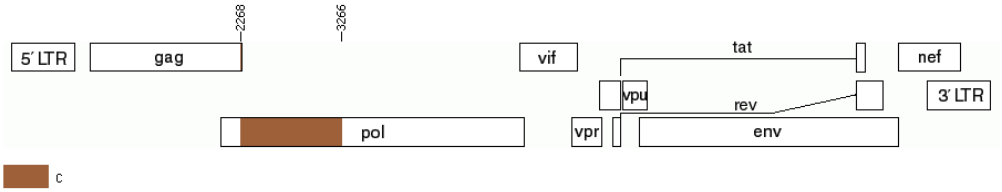
jpHMM result:

Sequence #1: **>Seq18**

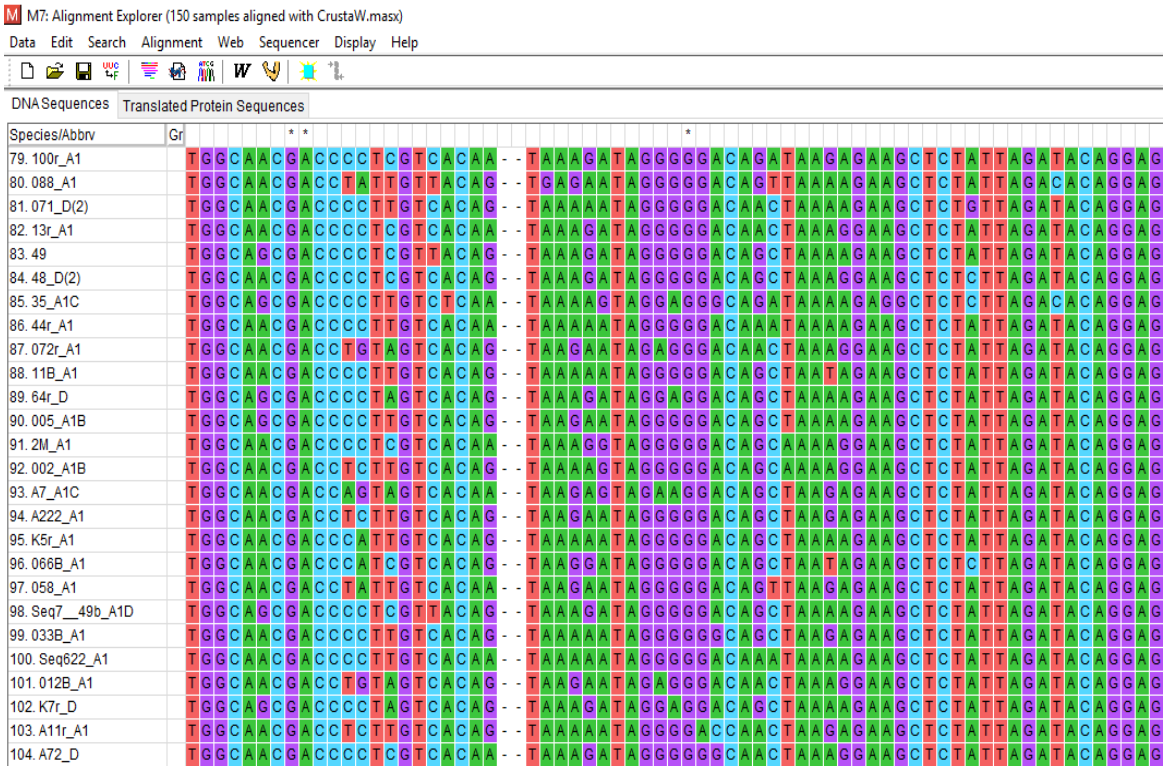
This sequence is related to subtype(s): **C**

Fragment Start Position	Uncertainty Region Start - End	Breakpoint Interval Start - End	Fragment End Position	Fragment Subtype
Position in the original sequence [pred_recombination], [recombination_incl_UR_and_BPI], [UR_and_BPI]				
1	-	-	**999**	**C**
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2268	-	-	**3266**	**C**

Genome map (based on HXB2 numbering)



Appendix 10:- Alignment of sequences by CrustalW in MEGA7


M7: Alignment Explorer (150 samples aligned with CrustaW.masx)
Data Edit Search Alignment Web Sequencer Display Help
DNA Sequences
Translated Protein Sequences
Species/Abbrv
Gr
79. 100r_A1
80. 088_A1
81. 071_D(2)
82. 13r_A1
83. 49
84. 48_D(2)
85. 35_A1C
86. 44r_A1
87. 072r_A1
88. 11B_A1
89. 64r_D
90. 005_A1B
91. 2M_A1
92. 002_A1B
93. A7_A1C
94. A222_A1
95. K5r_A1
96. 066B_A1
97. 058_A1
98. Seq7__49b_A1D
99. 033B_A1
100. Seq622_A1
101. 012B_A1
102. K7r_D
103. A11r_A1
104. A72_D


M7: Alignment Explorer (150 samples aligned with CrustaW.masx)
Data Edit Search Alignment Web Sequencer Display Help
DNA Sequences
Translated Protein Sequences
Species/Abbrv
Gr
105. A65_A1
106. A6r_D
107. A9r_A1
108. A4r_A1
109. A13_D
110. A1_A1
111. 468_A1
112. 467_A1
113. 465_A1
114. 456_A1B
115. 455_A1
116. 453_A1
117. 452_A1
118. 451_A1C
119. 406_A1
120. 390_A1B
121. 389_A2D
122. 386_D
123. 361_D
124. 359_D
125. 353_A1
126. 352_A1
127. 349_D
128. 321_A1B
129. 319_A1
130. 309_A1


M7: Alignment Explorer (150 samples aligned with CrustaW.masx)
Data Edit Search Alignment Web Sequencer Display Help
DNA Sequences
Translated Protein Sequences
Species/Abbrv
Gr
131. 304_D
132. 275_A1D
133. 268_A1
134. 266_A1K
135. 145_A1
136. 112j_A1
137. 082_A1A2
138. 071_D
139. 066_A1
140. 043B_A1
141. 049_A1J
142. 48_D
143. 090r_D
144. 047_A1B
145. 035_A1B
146. 034_A1A2
147. 03r_A1
148. B11_A1
149. 02r_A1
150. 010_A1
151. 4r-D

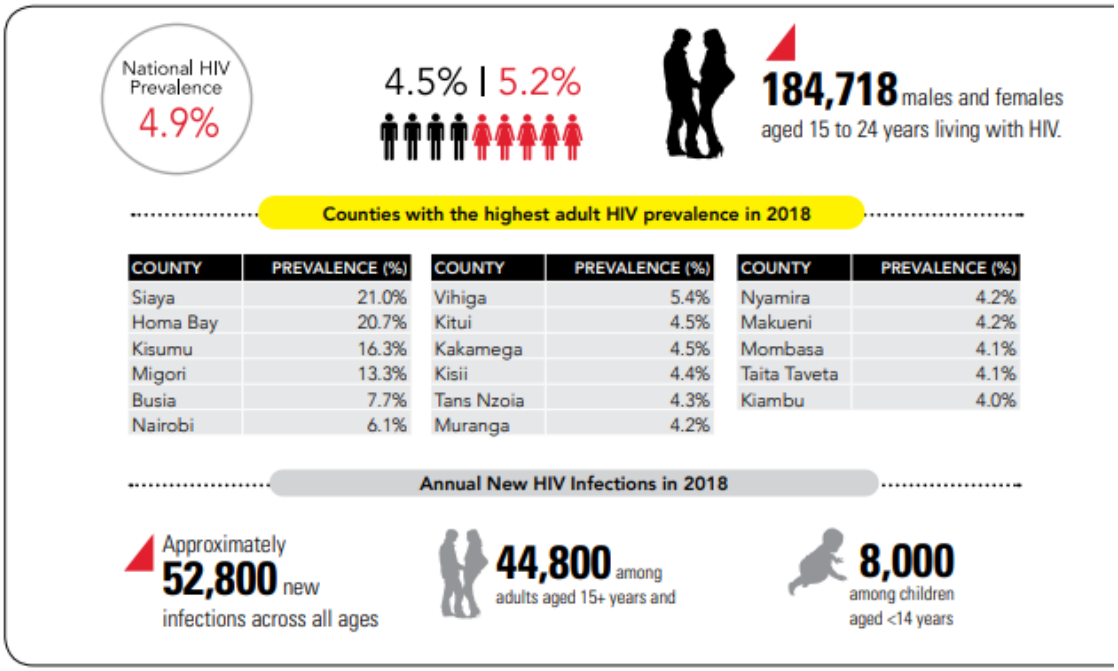
Appendix 11: - Questionnaire utilized to Collect of Data

Appendix 6: Laboratory form/Questionnaire used to collect data for analysis

Prevalence and molecular characterization of HIV-1 drug-resistance patterns among patients visiting selected health facilities in Busia county, Kenya

1. Date of collection.....MAY 2020.....
2. Name of Health Facility .....MATAYOS.....
3. Sample code .....A30.....
4. Sex .....F.....
5. Age of a patient .....63.....
6. Date of sample analysis.....JULY 2020.....
7. Type of ARVs the patient is taking ...TDF+3TC+ATV/r..Class.....2ND.....
8. For how long a patient has been on the above ARV? .....7 YEARS.....
9. Previous ARVs prior to the above...AS ABOVE.....Class.....AS ABOVE
10. Viral load (copies/ml) .....47,060 copies/ml of blood.....
11. HIV-1 subtype(s) .....B.....
12. Presence of HIV-1 drug resistance .....Yes....✓.....No.....
13. Mutation Type .....NNRTIS + NRTIS.....
    i) NNRTIS .....A98G, K101E, Y181C, G190A.....
    ii) NRTIs .....A62V, K65R.....
    iii) PIs .....MAJOR – NONE.....

Appendix 12: - Prevalence of HIV amongst Counties of Kenya



Source: Kenya HIV Estimates Report, 2018. https:// nacc. or. Ke /wp-content/ uploads/ 2018 /12/ HIV- estimates – report – Kenya - 20182. pdf