Organic Cation Transporter 2 (*OCT2/SLC22A2*) Gene Variation in the South African Bantu-Speaking Population and Functional Promoter Variants

Nina C. Wilson,¹ Ananyo Choudhury,² Nadia Carstens,³ and Demetra Mavri-Damelin¹

Abstract

SLC22A2 facilitates the transport of endogenous and exogenous cationic compounds. Many pharmacologically significant compounds are transported by *SLC22A2*, including the antidiabetic drug metformin, anticancer agent cisplatin, and antiretroviral lamivudine. Genetic polymorphisms in *SLC22A2* can modify the pharmacokinetic profiles of such important medicines and could therefore prove useful as precision medicine biomarkers. Since the frequency of *SLC22A2* polymorphisms varies among different ethnic populations, we evaluated these in South African Bantu speakers, a majority group in the South African population, who exhibit unique genetic diversity, and we subsequently functionally characterized promoter polymorphisms. We identified 11 polymorphisms within the promoter and 9 single-nucleotide polymorphisms (SNPs) within the coding region of *SLC22A2*. While some polymorphisms appeared with minor allele frequencies similar to other African and non-African populations, some differed considerably; this was especially notable for three missense polymorphisms. In addition, we functionally characterized two promoter polymorphisms; rs138765638, a three base-pair deletion that bioinformatics analysis suggested could alter c-Ets-1/2, Elk1, and/or STAT4 binding, and rs59695691, an SNP that could abolish TFII-I binding. Significantly higher luciferase reporter gene expression was found for rs138765638 (increase of 37%; $p=0.001$) and significantly lower expression for rs59695691 (decrease of 25%; $p=0.038$), in comparison to the wild-type control. These observations highlight the importance of identifying and functionally characterizing genetic variation in genes of pharmacological significance. Finally, our data for *SLC22A2* attest to the importance of considering genetic variation in different populations for drug safety, response, and global pharmacogenomics, through, for example, projects such as the Human Heredity and Health in Africa initiative.

Keywords: pharmacogenetics, personalized medicine, genetic variation, South Africa, organic cation transporters

Introduction

GENETIC POLYMORPHISMS within drug transporter genes have been shown to contribute to variation in drug response among individuals (Yee et al., 2010). Broad-specificity transporters, including the solute carrier superfamily 22 (*SLC22*), are responsible for the absorption, distribution, and excretion of various endogenous and exogenous substrates, and this family includes the organic anion transporters, the organic cation transporters (OCTs), and the organic cation and zwitterion transporters (Koepsell et al., 2007). Of specific relevance to this study is the OCT member OCT2, also known as

SLC22A2, which is encoded for by the 45 kilobase (kb) *SLC22A2* gene situated on chromosome 6q26 and comprises 11 coding exons (Gorboulev et al., 1997; Gründemann and Schömig, 2000; Koehler et al., 1997).

SLC22A2 expression not only occurs mainly at the basolateral membrane of renal proximal tubule cells in the kidney but is also found in the placenta, skin, spleen, lung, inner ear, small intestine, thymus, and brain (Koepsell et al., 2007). *SLC22A2* facilitates the transport of endogenous and exogenous cationic compounds (Table 1), with one major physiological function being the removal of these from the circulation into renal epithelial cells of the kidney, where

¹The School of Molecular and Cell Biology, University of the Witwatersrand, Johannesburg, South Africa.

²Sydney Brenner Institute for Molecular Bioscience (SBIMB), University of the Witwatersrand, Johannesburg, South Africa.

³Division of Human Genetics, National Health Laboratory Service, Johannesburg, South Africa.

TABLE 1. A SELECTION OF PHARMACOLOGICALLY SIGNIFICANT CATIONIC COMPOUNDS TRANSPORTED BY SLC22A2

Compound	Drug category	References
Acetylsalicylate	NSAID	Fujita et al. (2006)
Amantadine	Antiparkinsonian, antiviral	Busch et al. (1998)
Cisplatin	Anticancer	Yonezawa et al. (2006)
Lamivudine	Antiretroviral	Jung et al. (2008)
Memantine	Antiparkinsonian	Busch et al. (1998)
Metformin	Antidiabetic	Kimura et al. (2005)
Oxaliplatin	Anticancer	Yonezawa et al. (2006)
Quinine	Antimalarial	Gorboulev et al. (1997)
Salicylate	NSAID	Fujita et al. (2006)
Zalcitabine	Antiretroviral	Jung et al. (2008)

NSAID, nonsteroidal anti-inflammatory drug.

they are finally excreted into the urine (Motohashi et al., 2002). As such, SLC22A2 plays a significant role in the pharmacokinetic profiles of cationic drugs.

Genetic polymorphisms in the *SLC22A2* coding region have been identified and characterized in various ethnic populations and these have shown that indeed some coding region polymorphisms result in missense mutations that can significantly alter transporter function of SLC22A2 substrates (Fukushima-Uesaka et al., 2004; Leabman et al., 2002; Wang et al., 2007). In comparison, noncoding and promoter polymorphisms have been less well characterized and their biological effects are more difficult to ascertain (Ogasawara et al., 2008), but could still have a significant impact on SLC22A2 function.

While the genetic diversity among African populations is greater when compared to other populations, there are relatively few studies regarding genetic variation in drug transporter genes in South Africans, despite the fact that populations within South Africa have unique genetic profiles and allele frequencies, which can differ considerably from other African and non-African populations (Gurdasani et al., 2014; Ramsay, 2012; Tishkoff et al., 2009; Warnich et al., 2011). A major factor for these differences is the relatively high proportion of Khoisan admixture in these populations (Barbieri et al., 2014; Busby et al., 2016; Pickrell et al., 2012, 2014). Determining genetic polymorphisms with an aim to understanding their effect on transporter efficacy in the South African Bantu-speaking population is significant in the context of precision/personalized treatment.

The profile of drugs that SLC22A2 transports is very important for the South African population as these include the most commonly prescribed antidiabetic drug, metformin; the antimalarial, quinine; and the antiviral, lamivudine (Table 1). Therefore, identifying and characterizing genetic variation could have a significant impact on understanding patient response to these drugs.

In this study, we identified genetic polymorphisms within the *SLC22A2* promoter and coding regions in the majority Bantu-speaking South African population and subsequently functionally characterized two promoter polymorphisms.

Materials and Methods

Research ethics statement

Permission to carry out this study was granted from the Human Research Ethics Committee (Medical) of University

of the Witwatersrand, Johannesburg, #M10745 and #M120640. A written informed consent was obtained from all subjects.

Identification of genetic polymorphisms in the *SLC22A2* gene

Genetic polymorphisms were identified within *SLC22A2* (GRCh38: ENSG00000112499) promoter region (6:160260216–6:160258477) and coding regions (6:160216762–6:160259016), using publically available databases from various studies examining genetic variation within Bantu-speaking individuals. These included the whole-genome sequencing study of 40 Bantu-speaking individuals (personal correspondence with N. Carstens, 2016; unpublished data) and 100 Zulu individuals (Gurdasani et al., 2014).

We performed Sanger sequencing for the *SLC22A2* promoter on 10 Bantu-speaking individuals to isolate and subsequently characterize two promoter polymorphisms that we determined could be of functional significance. We considered the recent study by Jacobs et al. (2015) that evaluated specifically the single-nucleotide polymorphisms (SNPs) in *SLC22A2* in 96 Xhosa individuals. We compared minor allele frequency (MAF) values for known polymorphisms from the South African data to the overall MAF obtained from the 1000 Genomes Project (The 1000 Genomes Project Consortium, 2015). We compared these values to other African populations and to the overall MAF values excluding the African population, which were calculated to determine allele frequencies within the non-African populations (i.e., European, American, South Asian, and East Asian). These comparisons allowed us to determine how similar or dissimilar the South African Bantu-speaking population is to other non-African populations and to other African populations, since South Africa is not represented in these groups (Table 2).

Prediction of effect of polymorphism on expression and function

We screened the polymorphisms identified in the coding region using the Variant Effect Predictor tool (McLaren et al., 2016). Potential deleterious effects were predicted using the Sorting Intolerant From Tolerant (SIFT) and Polymorphism Phenotyping (Polyphen) programs through the Ensembl Genome Browser (Adzhubei et al., 2010; Ng and Henikoff, 2003). Modifications to potential transcription factor binding sites within the *SLC22A2* promoter were predicted using ALGGEN PROMO version 3.0.2 (Farré et al., 2003; Messager et al., 2002), through which the wild-type or mutant alleles with 15 base-pair flanking sequences were screened for whether the polymorphisms could modify putative transcription factor binding.

Functional characterization of the *SLC22A2* promoter polymorphisms

We performed Sanger sequencing on 10 Bantu-speaking individuals to isolate and functionally characterize the promoter polymorphisms rs59695691 and rs138765638, which bioinformatics analysis predicted to affect transcription factor binding. Ethics approval for this study was obtained from the Human Research Ethics Committee (Medical) of the University of the Witwatersrand, Johannesburg #M10745 and #M120640.

TABLE 2. THE STUDIES USED FOR ANALYSIS AND COMPARISON OF GENETIC POLYMORPHISMS IDENTIFIED IN THE *SLC22A2* PROMOTER AND CODING REGIONS IN THE SOUTH AFRICAN BANTU-SPEAKING POPULATION

Source	Population	Description	n	Type of sequencing
Personal correspondence with N. Carstens, 2016; unpublished data	Bantu speaking	South African	40	Exome (HC)
Gurdasani et al. (2014)	Zulu	South African	100	WGS (LC)
1000 Genomes Project—Phase 3	AFR	African	661	WGS (LC)
	ACB	African Caribbean in Barbados	96	
	ASW	African Ancestry in Southwest USA	61	
	ESN	Esan in Nigeria	99	
	LWK	Luhya in Webuye, Kenya	99	
	GWD	Gambian in Western Division, The Gambia	113	
	MSL	Mende in Sierra Leone	85	
	YRI	Yoruba in Ibadan, Nigeria	108	

HC, high coverage; LC, low coverage; n, number of subjects; WGS, whole-genome sequencing.

First, DNA was extracted from blood using a standard salting-out method, and then, quality was assessed using gel electrophoresis and quantified using a NanoDrop 1000 spectrophotometer. Next, we designed primers to amplify across a 1718 bp region of the *SLC22A2* promoter (−1421 to +277 relative to the translation start site). Polymerase chain reaction (PCR) products were sequenced three times in both the forward orientation and reverse orientation to confirm the presence of the polymorphisms (Inqaba Biotechnical Industries).

For cloning into the pGL4.10 luciferase reporter vector (Promega), these primers were redesigned to incorporate *NheI* and *EcoRV* restriction enzyme sites to allow for directional cloning into pGL4.10: forward primer, OCT2FNhe 5'-ATAAGCTAGCCCCCTGATGTGTGAGAGCAG-3' and reverse primer, OCT2REco 5'-GAGCGATATCCGTAGCGCCTACACTGTCTTG-3' (the underlined regions represent the restriction sites). The promoter region was amplified using Kapa HiFi high-fidelity DNA polymerase (Kapa Biosystems), with the following cycling conditions: 95°C for 3 min; 30 cycles of 98°C for 20 sec, 72°C for 15 sec, 72°C for 1.5 min; 72°C for 1.5 min; and a 4°C hold. The amplicon was restricted and subsequently ligated into the pGL4.10 vector using T4 DNA ligase (Kapa Biosystems), and competent JM109 strain of *E. coli* was transformed and selected on 100 µg/mL ampicillin-containing LB agar plates.

Recombinant vectors were identified through colony PCR with primers designed to amplify across the vector multiple cloning site pGL4.10FWD 5'-CTAGCAAAATAGGCTGTCCC-3' and pGL4.10REV 5'-TTCATGGCTTTGTGCAGCT-3' using Kapa Taq ReadyMix (Kapa Biosystems) with the following cycling conditions: 95°C for 2 min; 25 cycles of 95°C for 30 sec, 57°C for 30 sec, 72°C for 2 min; 72°C for 2 min, and a 4°C hold. A standard alkaline lysis plasmid extraction protocol was used to isolate the vector from transformed cells that were cultured for 16 h in LB broth supplemented with 100 µg/mL ampicillin. The three different *SLC22A2* promoter sequences (wild type, rs59695691, and rs138765638) were confirmed by sequencing and alignment with the reference sequence before analysis by luciferase assay. The wild-type promoter sequence used as the control is the one that appears in Ensembl for *SLC22A2* (ENSG00000112499; GRCh38; 6:160260216—6:160258477).

Next, the luciferase assay was performed to determine expression levels of luciferase from the wild type and SNP containing *SLC22A2* promoter vectors. The MRC-5 human fetal lung fibroblast cell line was maintained in a humidified atmosphere of 5% CO₂ and 95% air at 37°C and cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum with 1% penicillin–streptomycin. Cells were seeded into 96-well culture plates at a density of 12,000 cells per well. After 24 h, when the cells had reached ~80% confluency, cells were transfected with 200 ng of vector using the TurboFect transfection reagent (Thermo Fisher Scientific).

An empty (promoterless) pGL4.10 vector was used as a negative control. After 24 h, the plates were removed from the 37°C incubator, left to stand at room temperature for 15 min, and then 100 µL of the Steady-Glo[®] luciferase assay substrate (Promega) added to each well. After 5 min at room temperature, each well was mixed by pipetting and transferred into a white 96-well plate.

The expression level of luciferase, in the form of light, was measured using a GloMax 96 microplate luminometer (Promega). The expression of luciferase was determined for each vector and these were compared to the wild-type *SLC22A2* promoter. Statistical significance was analyzed on $n = 4$ using a paired *t*-test, and $p < 0.05$ considered to be statistically significant.

Results

Genetic polymorphisms within the *SLC22A2* promoter region

Overall, a total of 11 genetic polymorphisms were identified within the *SLC22A2* promoter region in the Bantu-speaking population (Table 3). We then screened each polymorphism, with a 15 bp flanking sequence on either side, using the ALGGEN PROMO program, which identifies potential alterations in transcription factor binding sites. In this way, we determined whether each promoter polymorphism may have the potential to alter gene expression levels. We identified two polymorphisms that may influence the binding of transcription factors to the DNA.

First, the SNP rs59695691, which is found at −195 relative to the translation start site, lies in the binding site for TFII-I, an integral member of the basal transcription machinery (Roy,

TABLE 3. GENETIC POLYMORPHISMS IDENTIFIED WITHIN THE *SLC22A2* PROMOTER

dbSNP ID	WT allele	Minor allele	MAF SA ^a		MAF AFR ^c	MAF All		Nucleotide position (GRCh38)	Nucleotide position (GRCh37)	Position from ATG ^e	Location
			Bantu (n=40)	MAF SA ^b Zulu (n=100)		excl. AFR	AFR ^d				
rs59695691	T	C	0.025	0.085	0.023	<0.01	0.006	160,258,952	160,679,984	-195	Upstream/5' UTR/intronic
rs5920607	G	A	0.063	0.09	0.045	<0.01	0.014	160,259,003	160,680,035	-246	Upstream/5' UTR/intronic
rs60249401	C	T	0.025	0.015	0.020	<0.01	0.007	160,259,181	160,680,213	-424	Upstream/intronic
rs138765638	CTT	—	—	0.21	0.138	0.084	0.1	160,259,707–160,259,709	160,680,739–160,680,741	-953 to -955	Upstream/5' UTR
rs113384645	C	T	0.032	0.045	—	—	—	160,259,735	160,680,767	-978	Upstream/5' UTR
rs80154852	G	A	0	0.095	0.057	<0.01	0.016	160,259,902	160,680,934	-1146	Upstream/5' UTR
rs146811048	G	A	0.10	0.03	0.023	<0.01	0.006	160,259,946	160,680,978	-1189	Upstream/5' UTR
rs148965379	A	T	0.091	0.03	0.023	<0.01	0.006	160,259,979	160,681,011	-1222	Upstream/5' UTR
rs66512417	—	TGAA	—	0.045	0.014	0.090	0.071	160,260,027–160,260,030	160,681,059–160,681,062	-1270 to -1273	Upstream/5' UTR
rs316023	T	C	0.261	0.16	0.305	0.469	0.420	160,260,282	160,681,314	-1525	Upstream/5' UTR
rs3127573	A	G	0.44	0.21	0.138	0.086	0.101	160,260,361	160,681,393	-1604	Upstream/5' UTR

^aBantu-speaking individuals-40 whole genome (personal communication with N. Carstens, 2016; unpublished data).^bZulu individuals (Gurdasani et al., 2014).^cAfrican populations from the 1000 Genomes Project Phase 3.^dAll individuals from the 1000 Genomes Project Phase 3: African, American, East Asian, European, and the South Asian.^eIndicates position relative to the translation start site.

—, either there is no information available or the variation could not be identified; MAF, minor allele frequency; n, number of individuals; SA, South African; SNP, single-nucleotide polymorphism; WT, wild type.

FIG. 1. Functional characterization of the effects of rs59695691 and rs138765638 on luciferase reporter gene expression. The rs59695691-containing vector, which carries the T>C base change at -195 (-195-SNP), showed a decrease in luciferase reporter expression (relative light units), while the rs138765638-containing vector with the CTT deletion showed an increase in expression, in comparison to the wild-type control (WT), $n=4 \pm \text{SD}$. SD, standard deviation; SNP, single-nucleotide polymorphism.

2001). Second, rs138765638, a three base-pair deletion, was predicted to modify transcription factor binding of c-Ets-1/2, Elk1, and/or STAT4. We then functionally analyzed these polymorphism promoters by cloning these promoter regions into luciferase reporter vectors. We observed a significant decrease of 25% in luciferase expression from the rs59695691-containing promoter pGL4.10 vector, and a significant increase of ~37% in luciferase expression for rs138765638, in comparison to the wild-type control (Fig. 1).

Genetic polymorphisms within the *SLC22A2* coding regions

A total of nine SNPs were identified within the *SLC22A2* coding region in the South African Bantu-speaking population (Table 4). Five SNPs were found to be synonymous with the rs772717144 (Val94Val) and rs58264151 (Gly370Gly) variants identified only in the South African Bantu-speaking populations, although at low frequency, and were not seen in any of the 1000 Genomes Project populations. The polymorphism rs624249 (Thr130Thr) was found at a lower frequency compared to the African and non-African populations, which exhibit similar MAF values, while rs316003 (Val502Val) showed a notably higher allele frequency in the African population in comparison to the South African populations in this study and to the non-African population. Four SNPs were identified as nonsynonymous, missense mutations, these being rs316019 (Ser270Ala), rs8177516 (Arg400Cys), rs8177517 (Lys432Gln), and rs139045661 (Ile552Asn). The MAF of Ser270Ala was found to be similar to all other populations. Interestingly, Arg400Cys, Lys432Gln, and Ile552Asn all appear in the South African populations in our study, and seem to be prevalent in African over non-African populations in general.

Discussion

Genetic polymorphisms within *SLC22A2* have been shown to change the expression level or activity of the protein and

TABLE 4. SINGLE-NUCLEOTIDE POLYMORPHISMS IDENTIFIED WITHIN THE *SLC22A2* CODING REGIONS

<i>dbSNP ID</i>	<i>WT allele</i>	<i>Minor allele</i>	<i>MAF SA^a</i> <i>Bantu</i> <i>(n = 40)</i>	<i>MAF SA^b</i> <i>Zulu</i> <i>(n = 100)</i>	<i>MAF AFR^c</i>	<i>MAF All excl. AFR</i>	<i>MAF All^d</i>	<i>Nucleotide position (GRCh38)</i>	<i>Nucleotide position (GRCh37)</i>	<i>Variant effect</i>	<i>Amino acid change</i>
rs772717144	C	T	0	0.005	—	—	—	160,258,476	160,679,508	Synonymous	Val94Val
rs624249	C	A	0.188	0.11	0.266	0.264	0.264	160,258,368	160,679,400	Synonymous	Thr130Thr
rs112210325	C	A	0	0	0.002	0	0.001	160,258,359	160,679,391	Synonymous	Ser133Ser
rs316019	C	A	0.20	0.09	0.185	0.118	0.137	160,249,250	160,670,282	Missense	Ser270Ala ^e
rs58264151	G	A	0.013	0.005	—	—	—	160,243,741	160,664,773	Synonymous	Gly370Gly
rs8177516	G	A	0.013	0.045	0.013	0	<0.01	160,243,653	160,664,685	Missense	Ala400Cys ^e
rs8177517	T	G	0.038	0.015	0.039	<0.01	0.010	160,242,388	160,663,420	Missense	Lys432Gln ^e
rs316003	C	T	0.40	0.425	0.576	0.207	0.307	160,224,800	160,645,832	Synonymous	Val502Val
rs139045661	A	T	0.063	0.015	0.012	0	<0.01	160,217,445	160,638,477	Missense	Ile552Asn

^aBantu-speaking individuals (personal communication with N. Carstens, 2016; unpublished data).

^bZulu individuals (Gurdasani et al., 2014).

^cAfrican populations within the 1000 Genomes Project Phase 3.

^dAll individuals from the 1000 Genomes Project Phase 3: African, American, East Asian, European, and the South Asian.

^eThese missense SNPs have been functionally characterized by Leabman et al. (2002).

—, there is no information available; *n*, number of individuals; WT, wild type.

these have been documented to modify the pharmacokinetics of *SLC22A2* substrates (Ieiri et al., 2006). Many pharmacologically significant drugs utilize *SLC22A2* for transport and these include the some of the most commonly prescribed antidiabetic, antiviral, and anticancer drugs, including quinine for malaria and lamivudine for HIV, which are of considerable importance for the South African population (Blumberg, 2015; Kimura et al., 2005; Lazarus et al., 2013). Due to the unique level of genetic diversity within the African continent, along with advancement in technologies that allow for personalized medicine, there has been an increased drive to identify and characterize genetic variation in African populations (Ramsay, 2012).

Recent large-scale genome-sequencing projects are providing a more comprehensive view of African genomic diversity; the African Genome Variation Project has generated 300 whole-genome sequences from Eastern and Southern African populations (Gurdasani et al., 2014), and more such data are expected from ongoing projects such as the Southern African Human Genome Programme, the Human Heredity and Health in Africa (H3Africa) Project, and the Trypano-GEN study. We focused on the identification of genetic polymorphisms in the *SLC22A2* promoter and coding regions in the South African Bantu-speaking population, as these are yet to be identified and characterized, especially since this gene seems to exhibit considerable population-specific genetic variation that could impact on protein function and hence patient drug response.

Polymorphisms in the promoter region of a gene can have a significant impact on gene expression levels through various mechanisms, including the modification of transcription factor binding sites. We identified and subsequently functionally characterized key promoter polymorphisms. We found that 6 of the 11 polymorphisms identified had frequencies that ranged from 1.4% up to 9.5% in the South African Bantu-speaking populations in this study, while in non-African populations these appeared at less than 1%. In comparison to the African population, some of these were similar but some were dissimilar.

We also found that SNP rs113384645 was observed only in the South African population and not in any of the 1000 Genomes Project populations. Of the remaining four polymorphisms that displayed a MAF of greater than 0.01 in the

non-African population, the polymorphisms rs138765638 and rs3127573 displayed considerably lower MAF in non-Africans, in comparison to the African and South African Bantu-speaking populations, whereas the polymorphisms rs66512417 and rs316023 were notably lower MAFs in the African and South African populations, in comparison to the non-African populations.

We then considered the potential functional effect of two of the promoter polymorphisms, rs59695691 and rs138765638. We focused on these since bioinformatics analysis suggested that the transcription factor binding sites could be modified by these two polymorphisms. The SNP rs59695691 lies in the binding site for TFII-I, member of the basal transcription machinery that assembles around the core promoter. TFII-I also acts as an activator that binds to upstream regulatory regions and thereby facilitates communication between these upstream transcriptional regulatory sites and the core promoter (Roy, 2001). ALGGEN PROMO predicted that the TFII-I site 5'-GGACA/GT-3' modified by the A>G SNP, as indicated, could abolish TFII-I binding.

We observed these polymorphisms in the South African populations with an MAF 0.085 for the Zulu-only study and 0.025 in the South African Bantu study, while Jacobs et al. (2015) identified this SNP with an MAF of 0.263 in the Xhosa population (*n* = 96). In other African and non-African populations, rs59695691 was observed at 0.023 and <0.01, respectively. As such, we decided to functionally characterize this SNP using the luciferase reporter assay, considering the fundamental role TFII-I has in transcription. As anticipated, luciferase reporter gene expression decreased by a significant 25% from this SNP-containing promoter region.

We also analyzed rs138765638, a 3-bp deletion, since ALGGEN PROMO predicted that the binding of the transcription factors c-Ets-1/2, Elk1, and/or STAT4 could be affected. While there is relatively little characterization of the transcription factors that regulate *SLC22A2* expression, STAT4 and *SLC22A2* have overlapping expression patterns, such as in parts of the brain. We found that this deletion polymorphism, which was identified with an MAF of 0.21 in the South African Zulu population, 0.138 in Africans, and at 0.084 in non-Africans, showed a significant 37% increase in luciferase reporter gene expression.

While the physiological effects of promoter polymorphisms can be difficult to ascertain, they have previously been shown to affect *SLC22A2* gene expression, and as such may affect individuals carrying these gene variants. Ogasawara et al. (2008) identified a functionally significant genetic polymorphism in the *SLC22A2* promoter in 63 Japanese nephrectomized patients. An AAG deletion at position –578 to –576 (–578_–576delAAG) was shown to significantly reduce the activity of the *SLC22A2* promoter and heterozygotes of this deletion were shown to have lower *SLC22A2* mRNA levels.

With regard to coding region polymorphisms, we identified a number of synonymous mutations with high MAF values, as well as being considerably different to other African and non-African populations. While it is difficult to determine the role of such synonymous genetic variations, their effect on *SLC22A2* cannot be discounted as they may alter protein expression and function through pretranscriptional, posttranscriptional, and posttranslational modifications, thereby influencing patient drug response (Sauna and Kimchi-Sarfaty, 2011).

We also identified four missense mutations in the coding region of *SLC22A2*. The MAF of rs316019 (Ser270Ala) was found to be similar to all other populations, and since it is found in up to 14% of the population, it has been the subject of a number of functional characterization studies (Leabman et al., 2002; Tzvetkov et al., 2009; Zolk, 2009). However, whether this amino acid change alters the function of *SLC22A2* seems to depend on the specific cationic compound tested. For example, the renal clearance of metformin, which is transported by *SLC22A2*, seems to also be contradictory in various studies, with some reporting higher and others lower renal clearance, although study design and population ethnicity may play a significant role in these conflicting results (Chen et al., 2009; Song et al., 2008).

In addition, while mutations tend to be seen as negative traits, this Ser270Ala variant has also recently been associated with protective roles against cisplatin-induced nephro- and ototoxicity, which are serious side effects of cisplatin chemotherapy (Filipski et al., 2009; Lanvers-Kaminsky et al., 2015). For rs8177516 (Arg400Cys) and rs8177517 (Lys432Gln), which do not appear in non-African populations, we found these at between 1.3% and 4.5% in the African and South African populations. Previous functional characterization of these SNPs has shown that they deleteriously affect transporter function (Leabman et al., 2002). Finally, for rs139045661 (Ile552Asn), which also has a prevalence of less than 1%, there has not been any functional characterization as yet, although SIFT and PolyPhen programs predict this amino acid change to be tolerated.

In the context of precision medicine, our results could have a considerable impact on patient treatment in South Africa as well as for global science of rational therapeutics. Our data for *SLC22A2* raise the issues of considering genetic variation in different populations for the purposes of assessing drug safety and for drug response biomarkers, which further highlights the importance of projects such as the H3Africa initiative.

The role that pharmacogenetics will play in future global health must consider many factors, not least, the best course of clinical implementation, especially in resource-poor settings (Ozdemir et al., 2015). In addition, social and political science research on the ways in which precision medicine is currently emerging in resource-limited settings, shaping the local com-

munities, and the dependency of knowledge-based innovations on the context, power, and funding warrant further attention and interdisciplinary scholarship (Birch and Tyfield, 2013; De Vries, 2004; Petersen, 2013). With the continuing rise of precision medicine, we envisage that studies such as ours will enable patients in developing countries to benefit from specifically targeted treatment regimens.

Conclusions and Expert Outlook

Our study highlights differences and similarities between the South African Bantu-speaking population and other African and non-African populations. This was evident with the promoter polymorphisms that we functionally characterized and the missense mutations that appear at frequencies that are considerably different to other populations. In the age of precision medicine and considering the important role of *SLC22A2* in transporting drugs of considerable clinical significance, our finding warrants further investigation into the possible *in vivo* effects, especially during drug therapy where these polymorphisms may have additional confounding effects (Jung et al., 2008).

Furthermore, while many studies look at the pharmacokinetics of exogenous compounds, the effects of altered transporter function on endogenous substrates must also be taken into account, where, for example, decreased clearance of the neurotransmitters, norepinephrine or serotonin via *SLC22A2*, may lead to changes in mood-related behavior (Bacq et al., 2012).

In all, our research reinforces the fact that the genetic diversity in South Africans is distinct from other African populations and contributes to understanding genetic variation in *SLC22A2* in South African populations, which has important consequences for rational drug prescription and healthcare provided in South Africa as well as globally.

Acknowledgments

This study was funded by the University of the Witwatersrand, Johannesburg, and the National Research Foundation of South Africa with a grant to DMD (Grant No. 90710). NCW was supported by an NRF Masters Scholarship. A.C. was supported by the National Human Genome Research Institute of the National Institutes of Health under Award Number U54HG006938 (H3Africa Consortium). Any opinion, findings and conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect those of the authors' affiliated institutions or the funders.

Author Disclosure Statement

The authors declare that no conflicting financial interests exist.

References

- Adzhubei IA, Schmidt S, Peshkin L, et al. (2010). A method and server for predicting damaging missense mutations. *Nat Methods* 7, 248–249.
- Bacq A, Balasse L, Biala G, et al. (2012). Organic cation transporter 2 controls brain norepinephrine and serotonin clearance and antidepressant response. *Mol Psychiatry* 17, 926–938.

- Barbieri C, Güldemann T, Naumann C, et al. (2014). Unraveling the complex maternal history of Southern African Khoisan populations. *Am J Phys Anthropol* 153, 435–448.
- Blumberg LH. (2015). Recommendations for the treatment and prevention of malaria: Update for the 2015 season in South Africa. *S Afr Med J* 105, 175–178.
- Birch K, and Tyfield D. (2013). Theorizing the bioeconomy: Biovalue, biocapital, bioeconomics or...what? *Sci Technol Human Values* 38, 299–327.
- Busby GBJ, Band G, Le QS, et al. (2016). Admixture into and within sub-Saharan Africa. *eLife* 5, e15266.
- Busch AE, Karbach U, Miska D, et al. (1998). Human neurons express the polyspecific cation transporter hOCT2, which translocates monoamine neurotransmitters, amantadine, and memantine. *Mol Pharmacol* 54, 342–352.
- Chen Y, Li S, Brown C, Cheatham S, Castro RA, and Leabman MK. (2009). Effect of genetic variation in the organic cation transporter 2, OCT2, on the renal elimination of metformin. *Pharmacogenet Genomics* 19, 497–504.
- De Vries R. (2004). How can we help? From 'sociology in' bioethics to 'sociology of' bioethics. *J Law Med Ethics* 32, 279–292.
- Farré D, Roset R, Huerta M, et al. (2003). Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. *Nucleic Acids Res* 31, 3651–3653.
- Filipski KK, Mathijssen RH, Mikkelsen TS, Schinkel AH, and Sparreboom A. (2009). Contribution of organic cation transporter 2 (OCT2) to cisplatin-induced nephrotoxicity. *Clin Pharmacol Ther* 86, 396–402.
- Fujita T, Urban TJ, Leabman MK, Fujita K, and Giacomini M. (2006). Transport of drugs in the kidney by the human organic cation transporter OCT2 and its genetic variation. *J Pharm Sci* 95, 25–36.
- Fukushima-Uesaka H, Maekawa K, Ozawa S, et al. (2004). SNP communication fourteen novel single nucleotide polymorphisms in the SLC22A2 gene encoding human organic cation transporter (OCT2). *Drug Metab Pharmacokinet* 19, 239–244.
- Gorboulev V, Ulzheimer JC, Akhoundova A, et al. (1997). Cloning and characterization of two human polyspecific organic cation transporters. *DNA Cell Biol* 16, 871–881.
- Gründemann D and Schömig E. (2000). Gene structures of the human non-neuronal monoamine transporters EMT and OCT2. *Hum Genet* 106, 627–635.
- Gurdasani D, Carstensen T, Tekola-Ayele F, et al. (2014). The African Genome Variation Project shapes medical genetics in Africa. *North* 517, 327–332.
- Ieiri I, Takane H, Hirota T, Otsubo K, and Higuchi S. (2006). Genetic polymorphisms of drug transporters: Pharmacokinetic and pharmacodynamic consequences in pharmacotherapy. *Expert Opin Drug Metab Toxicol* 2, 651–674.
- Jacobs C, Pearce B, Du Plessis M, Hoosain N, and Benjeddou M. (2015). Single nucleotide polymorphisms of the SLC22A2 gene within the Xhosa population of South Africa. *Drug Metab Pharmacokinet* 30, 457–460.
- Jung N, Lehmann C, Rubbert A, et al. (2008). Relevance of the organic cation transporters 1 and 2 for antiretroviral drug therapy in human immunodeficiency virus infection. *Drug Metab Dispos* 36, 1616–1623.
- Kimura N, Masuda S, Tanihara Y, et al. (2005). Metformin is a superior substrate for renal organic cation transporter OCT2 rather than hepatic OCT1. *Drug Metab Pharmacokinet* 20, 379–386.
- Koehler MR, Wissinger B, Gorboulev V, Koepsell H, and Schmid M. (1997). The two human organic cation transporter genes SLC22A1 and SLC22A2 are located on chromosome 6q26. *Cytogenet Cell Genet* 79, 198–200.
- Koepsell H, Lips K, and Volk C. (2007). Polyspecific organic cation transporters: Structure, function, physiological roles, and biopharmaceutical implications. *Pharmacological Res* 24, 1227–1251.
- Lanvers-Kaminsky C, Sprowl JA, Malath I, et al. (2015). Human OCT2 variant c.808G>T confers protection effect against cisplatin-induced ototoxicity. *Pharmacogenomics* 16, 323–332.
- Lazarus EM, Otjombe K, Fairlie L, et al. (2013). Lamivudine monotherapy as a holding strategy in HIV-infected children in South Africa. *J AIDS Clin Res* 4, 246.
- Leabman MK, Huang CC, Kawamoto M, et al. (2002). Polymorphisms in a human kidney xenobiotic transporter, OCT2, exhibit altered function. *Pharmacogenet Genomics* 12, 395–405.
- McLaren W, Gil L, Hunt SE, et al. (2016). The ensembl variant effect predictor. *Genome Biol* 17, 122.
- Messeguer X, Escudero R, Farré D, Núñez O, Martínez J, and Mar Albà M. (2002). PROMO: Detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics* 18, 333–334.
- Motohashi H, Sakurai Y, Saito H, et al. (2002). Gene expression levels and immunolocalization of organic ion transporters in the human kidney. *J Am Soc Nephrol* 13, 866–874.
- Ng PC, and Henikoff S. (2003). SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res* 31, 3812–3814.
- Ogasawara K, Terada T, Motohashi H, et al. (2008). Analysis of regulatory polymorphisms in organic ion transporter genes (SLC22A) in the kidney. *J Hum Genet* 53, 607–614.
- Ozdemir V, Burke W, KHoury MJ, Knoppers BM, and Zimmer R. (2015). Genomics and public health. In: *Oxford Textbook of Global Public Health*. Detels R, Gulliford M, Karim QA, and Tan CC, eds. 6th ed. Oxford, UK: Oxford University Press, 140–161.
- Petersen A. (2013). From bioethics to a sociology of bio-knowledge. *Soc Sci Med* 98, 264–270.
- Pickrell JK, Patterson N, Barbieri C, et al. (2012). The genetic prehistory of southern Africa. *Nat Commun* 3, 1143.
- Pickrell JK, Patterson N, Lohc P-R, et al. (2014). Ancient west Eurasian ancestry in southern and eastern Africa. *Proc Natl Acad Sci U S A* 111, 2632–2637.
- Ramsay, M. (2012). Africa: Continent of genome contrasts with implications for biomedical research and health. *FEBS Lett* 586, 2813–2819.
- Roy AL. (2001). Biochemistry and biology of the inducible multifunctional transcription factor TFII-I. *Gene* 274, 1–13.
- Sauna ZE, and Kimchi-Sarfaty C. (2011). Understanding the contribution of synonymous mutations to human disease. *Nat Rev Genet* 12, 683–691.
- Song IS, Shin HJ, Shim EJ, et al. (2008). Genetic variants of the organic cation transporter 2 influence the disposition of metformin. *Clin Pharmacol Ther* 84, 559–562.
- The 1000 Genomes Project Consortium. (2015). A global reference for human genetic variation. *Nature* 526, 68–74.
- Tishkoff SA, Reed FA, Friedlaender FR, et al. (2009). The genetic structure and history of Africans and African Americans. *Science* 324, 1035–1044.
- Tzvetkov MV, Vormfelde SV, Balen D, et al. (2009). The effects of genetic polymorphisms in the organic cation transporters OCT1, OCT2, and OCT3 on the renal clearance of metformin. *Clin Pharmacol Ther* 86, 299–306.

- Wang Z-J, Yin OQP, Tomlinson B, and Chow MSS. (2007). OCT2 polymorphisms and in-vivo renal functional consequence: Studies with metformin and cimetidine. *Pharmacogenet Genomics* 18, 637–645.
- Warnich L, Drögemöller BI, Pepper MS, Dandara C, and Wright GE. (2011). Pharmacogenomic research in South Africa: Lessons learned and future opportunities in the rainbow nation. *Curr Pharmacogenomics Person Med* 9, 191.
- Yee SW, Chen L, and Giacomini KM. (2010). Pharmacogenomics of membrane transporters: Past, present and future. *Pharmacogenomics* 11, 475–479.
- Yonezawa A, Masuda S, Yokoo S, Katsura T, and Inui K. (2006). Cisplatin and oxaliplatin, but not carboplatin and nedaplatin, are substrates for human organic cation transporters (SLC22A1–3 and multidrug and toxin extrusion family). *J Pharm Exp Ther* 319, 879–886.
- Zolk O, Solbach TF, König J, and Fromm MF. (2009). Functional characterization of the human organic cation transporter 2 variant p.270Ala>Ser. *Drug Metab Dispos* 37, 1312–1318.

Address correspondence to:
Demetra Mavri-Damelin, PhD
The School of Molecular and Cell Biology
University of the Witwatersrand
Private Bag X3
Johannesburg WITS 2050
South Africa

E-mail: demetra.mavri-damelin@wits.ac.za

Abbreviations Used

H3Africa = Human Heredity and Health in Africa
MAF = minor allele frequency
OCT2 = organic cation transporter 2
SIFT = Sorting Intolerant From Tolerant
SLC = solute carrier
SNP = single-nucleotide polymorphism