Title

Differential expression of nuclear genes encoding mitochondrial proteins from urban and rural populations in Morocco

Journal

Cell Stress and Chaperones, 25(6)

ISSN

1355-8145

Authors

Bickler, Stephen W

Prieto, James M

Cauvi, David M

et al.

Publication Date

2020-11-01

DOI

10.1007/s12192-020-01108-x

Peer reviewed

Cell Stress and Chaperones (2020) 25:847–856
https://doi.org/10.1007/s12192-020-01108-x

ORIGINAL PAPER

Differential expression of nuclear genes encoding mitochondrial proteins from urban and rural populations in Morocco

Stephen W. Bickler[1,2,3] · James M. Prieto[4] · David M. Cauvi[2,3] · Victor De Cos[1] · Chanond Nasamran[5] · Emmanuel Ameh[6] · Said Amin[7] · Sneha Nicholson[1] · Hena Din[1] · Ana Olga Mocumbi[8,9] · Emilia Virginia Noormahomed[9] · Guillermo Tellez-Isaias[10] · Kathleen M. Fisch[5] · Antonio De Maio[2,3,11]

Received: 13 January 2020 / Revised: 26 March 2020 / Accepted: 30 March 2020 / Published online: 21 April 2020

Abstract

Urbanization in low-income countries represents an important inflection point in the epidemiology of disease, with rural populations experiencing high rates of chronic and recurrent infections and urban populations displaying a profile of noncommunicable diseases. To investigate if urbanization alters the expression of genes encoding mitochondrial proteins, we queried gene microarray data from rural and urban populations living in Morocco (GSE17065). The R Bioconductor packages edgeR and limma were used to identify genes with different expression. The experimental design was modeled upon location and sex. Nuclear genes encoding mitochondrial proteins were identified from the MitoCarta2.0 database. Of the 1158 genes listed in the MitoCarta2.0 database, 847 genes (73%) were available for analysis in the Moroccan dataset. The urban-rural comparison with the greatest environmental differences showed that 76.5% of the MitoCarta2.0 genes were differentially expressed, with 97% of the genes having an increased expression in the urban area. Enrichment analysis revealed 367 significantly enriched pathways (adjusted p value < 0.05), with oxidative phosphorylation, insulin secretion and glucose regulations (adj.p values = 6.93E-16) being the top three. Four significantly perturbed KEGG disease pathways were associated with urbanization—three degenerative neurological diseases (Huntington's, Alzheimer's, and Parkinson's diseases) and herpes simplex infection (false discover rate corrected p value (PGFdr) < 0.2). Mitochondrial RNA metabolic processing and translational elongation were the biological processes that had the greatest enrichment (enrichment ratios 14.0 and 14.8, respectively, FDR < 0.5). Our study links urbanization in Morocco with changes in the expression of the nuclear genes encoding mitochondrial proteins.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12192-020-01108-x) contains supplementary material, which is available to authorized users.

✉ Stephen W. Bickler
sbickler@health.ucsd.edu

James M. Prieto
jprieto4570@gmail.com

David M. Cauvi
dcauvi@health.ucsd.edu

Victor De Cos
Victor.decos@gmail.com

Chanond Nasamran
cnasamran@ucsd.edu

Emmanuel Ameh
eaameh@yahoo.co.uk

Said Amin
saidmamin@gmail.com

Sneha Nicholson
snicholson@rchsd.org

Hena Din
hdin@rchsd.org

Ana Olga Mocumbi
amocumbi@gmail.com

Emilia Virginia Noormahomed
enoormahomed@gmail.com

Guillermo Tellez-Isaias
tellez@uark.edu

Kathleen M. Fisch
kfisch@ucsd.edu

Antonio De Maio
ademaio@health.ucsd.edu

Extended author information available on the last page of the article

Keywords Humans · Urbanization · Noncommunicable diseases · RNA · Mitochondrial · Morocco · Gene expression

Introduction

Fifty-five percent of the world's population currently lives in a city, with this figure expected to rise to over two-thirds by 2050 (United Nations 2019). Urbanization is occurring most rapidly in Sub-Saharan Africa and Asia, where economic development and the population growth have been the fastest. Agricultural instability, forced migration due to political instability, ecological disasters (e.g., floods and climate change), and greater social and education opportunities in cities are other factors that lead to urbanization worldwide (Chen et al. 2014; Satterthwaite et al. 2010).

Changing environmental exposures are the sine qua non of urbanization. Low-income countries are perhaps the best place to observe these changes as urbanization is occurring rapidly and there are often extremes in levels of development between rural and urban areas. These changes include, but are not limited to, differences in rates of infection, diet, population density, nigh-time lighting, and levels of noise and pollution (Bickler 2000). Importantly, these different environmental exposures are associated with major changes in the disease epidemiology. In contrast to rural areas where chronic and recurrent infections predominate, urban populations often display a profile of noncommunicable diseases (NCDs), similar to that observed in high-income countries (Caldwell 2001; Eckert and Kohler 2014; Gong et al. 2012; Harpham 2009). This epidemiological shift from communicable to NCDs is a microcosm of that occurring globally for the past century (Jamison and Mosley 1991; Murray and Lopez 1997). NCDs are the leading cause of death in middle- and high-income countries. In 2015, 71% of the 58 million deaths occurring worldwide were related to NCDs, mainly cardiovascular diseases, cancer, and respiratory diseases (Mortality and Causes of Death 2016).

The biological mechanisms that underlie the shift from communicable to NCDs and their relationship to specific environmental factors remain poorly understood. To date, most research has focused on the impact of urbanization on cardiovascular risk factors such as serum lipids (Htet et al. 2017; Kavishe et al. 2019; Vasunilashorn et al. 2010), changes in the gut microbiome (Ayeni et al. 2018; Jha et al. 2018; Lokmer et al. 2020), or effects of pollution in urban areas (Amegah and Agyei-Mensah 2017; Coker and Kizito 2018). A less utilized approach has been to compare gene expression patterns in genetically similar rural and urban populations (Idaghdour et al. 2010, 2008; Nath et al. 2012). These gene expression studies have demonstrated the complex molecule changes that occur during the transition from a rural to urban environment. In a study from southern Morocco, the expression of over a third of the peripheral blood transcriptome was shown to differ between residents of a rural Berber village and the city of Agadir (Idaghdour et al. 2008). Similar results were observed in a larger follow-up study from Morocco (Idaghdour et al. 2010), and in an investigation on the peripheral blood transcriptomes of Fijians living in a rural village and the capital city of Suva (Nath et al. 2012). In all of these three studies, differences in the genome-wide expression signature were attributed to a combination of lifestyle, geography, and biotic factors. While these studies have been important in advancing our understanding of how environment shapes human biology, the broad changes in gene expression have been difficult to interpret, especially as they relate to specific risk factors and actual mechanisms.

An evolving application of these rural-urban datasets is to test specific hypotheses regarding how individual genes or pathways change with urbanization (Bickler et al. 2016, 2015, 2018). As an example, we used Moroccan gene expression database to show that urbanization alters the expression of G protein subunit genes, suggesting the possibility of environmentally specific G protein-coupled receptor (GPCR) signaling. Three genes controlling the phosphatidylinositol signaling pathway and one gene regulating cAMP (3'-5'-cyclic adenosine monophosphate) were significantly increased in the urban population. Further, the gene encoding the β-arrestin 1 protein (ARRB1), which dampens cellular responses to hormones, neurotransmitters, and other sensory signals (Buchanan and DuBois 2006) was increased in the rural population. This example illustrates the tremendous potential that exists for using molecular biological approaches to decipher the complex biological changes that occur during the transition from a rural to an urban environment.

In this investigation, we build on our previous efforts to understand the biological changes that occur with urbanization by testing the hypothesis that urbanization could alter mitochondrial gene expression. The rationale for investigating mitochondrial gene expression was twofold. First, urbanization was previously shown to enrich oxidative phosphorylation (Idaghdour et al. 2010)—a mitochondrial function. Second, mitochondria represent a cellular hub that connects energy metabolism, stress sensing, signaling, and cell survival (Lane 2005; Monlun et al. 2017; Naquet et al. 2016).

Material and methods

We queried gene microarray data from rural and urban populations living in Morocco for nuclear genes encoding proteins

associated with mitochondria (Fig. 1). We first identified differentially expressed genes in rural and urban populations and then selected out genes encoding mitochondrial proteins listed in the MitoCarta2.0 database. The differentially expressed MitoCarta2.0 database genes (i.e., the nuclear genes encoding mitochondrial proteins) were then analyzed using Gene Set Enrichment Analysis (GSEA), Signaling Pathway Impact Analysis (SPIA), and WebGestalt, and used to construct a gene interaction network.

Moroccan gene expression microarray data

Our analysis was done using gene expression microarray data (GSE17065) from the National Center for Biotechnology Information (NCBI, Bethesda, MD, USA). This dataset contains peripheral blood leukocyte gene expression data from individuals living in the Souss region of southern Morocco (Idaghdour et al. 2010). The Souss region is home to several million people of two dominant ethnicities living either in cities, or rural villages. One-half of the samples were from the high density, low- to middle-income city of Agadir. The remaining samples were from two rural villages, Boutroch which is predominantly Amazigh and quite isolated, and Ighrem, predominantly Arab with many of the men commuting to cities. As our desire was to better understand the gene expression differences in the most disparate urban-rural environments, we selected the Agadir-Boutroch comparison for our enrichment analysis.

Mitocarta2.0 genes

Nuclear genes encoding proteins with strong support of mitochondrial localization were identified from the MitoCarta2.0 database (https://www.broadinstitute.org/files/shared/metabolism/mitocarta/human.mitocarta2.0.html). This list of 1158 genes is based on mass spectrometry of mitochondria isolated from fourteen tissues, assessed protein localization through large-scale GFP tagging/microscopy, and six other genome-scale datasets of mitochondrial localization (Calvo et al. 2016).

Identification of differentially expressed genes in Moroccan urban and rural populations

Normalized microarray data for GSE17065 was downloaded from the NCBI using GEOquery package (Davis and Meltzer 2007). The R Bioconductor packages edgeR and limma (Ritchie et al. 2015) were used to implement the limma method for differential expression analysis (https://www.bioconductor.org). The experimental design was modeled upon location and sex treatment ($\sim 0 + \text{location} + \text{sex}$). Transcripts were annotated with the IllumnaHuman3.db (Dunning and Eldridge 2015), and filtered using genefilter (Gentleman et al. 2018). Similar to the analysis performed by Idaghdour et al. (Idaghdour et al. 2010), comparisons between multiple geographical locations such as Agadir-Bourtroch (urban-rural), Agadir-Ighrem (urban-rural), and Ighrem-Boutroch (rural-rural) were made. In addition, we estimated the average number of differentially expressed genes in the urban-rural populations by comparing Agadir to the average of the Boutroch and Ighrem groups. Significance was defined by using an adjusted cutoff of $p < 0.05$ after multiple testing corrections with a moderated t statistic in limma. In some cases, multiple illumina probes mapped to the same gene. The probe with the lowest p value was retained and all other probes were discarded.

Enrichment analysis of differentially expressed MitoCarta2.0 genes

Enrichment of the differentially expressed MitoCarta2.0 genes was assessed using Gene Set Enrichment Analysis (GSEA)

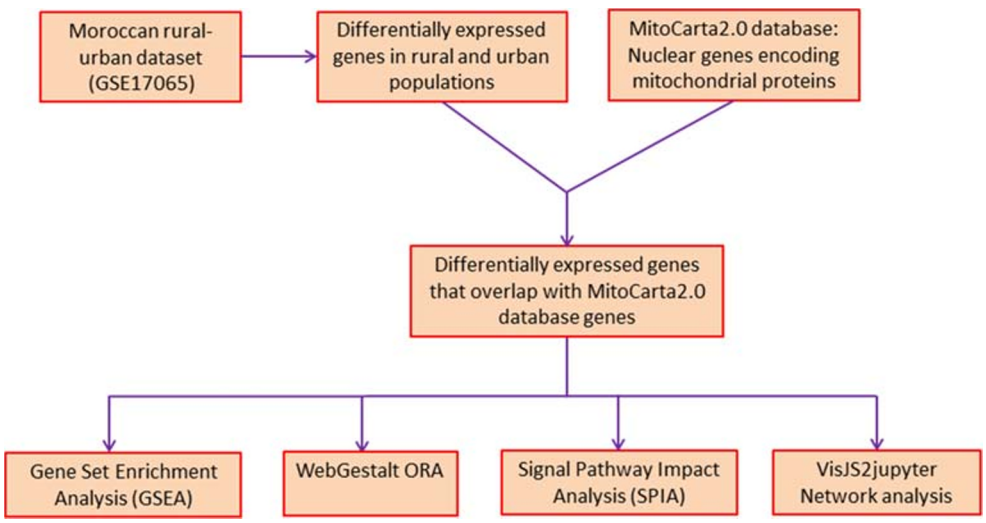

Fig. 1 Strategy for investigating if urbanization alters the expression nuclear genes encoding mitochondrial proteins

implemented in the Bioconductor GSVA package (Hanzelmann et al. 2013), over-representation analysis (ORA) in WEB-based Gene Set AnaLysis Tookit (WebGestalt)(Wang et al. 2017), and Signaling Pathway Impact Analysis (SPIA) (Tarca and Draghici 2018). GSEA assesses the enrichment of functionally related gene sets based on the Molecular Signatures Database (MSigDB) (Liberzon et al. 2015), while SPIA assesses the differentially expressed genes that change together in KEGG disease signaling pathways (Kanehisa et al. 2017). For the WebGestalt ORA, we used the gene ontology biological process database (Biological_ process_noRedundant) and the Ilumina_humant_12_v3 reference gene list.

Gene interaction network of differentially expressed MitoCarta2.0 genes

Significantly differentially expressed genes were used as seeds for network propagation (Cowen et al. 2017) on the STRING high confidence interactome (Szklarczyk et al. 2015). A graph-based modularity maximization clustering algorithm was used to identify groups of genes within the most proximal genes which were highly interconnected. Genes in the entire network and within each of these clusters were annotated with associated pathways identified by functional enrichment analysis, with the ToppGene (Chen et al. 2009). Network visualization and propagation were performed using Cytoscape (Shannon et al. 2003) and VisJS2jupyter (Rosenthal et al. 2018). The subgraph composed of the most proximal genes is visualized using a modified spring-embedded layout algorithm, modified by cluster membership, so that genes belonging to the same cluster are separated from other clusters. Differential expression log fold change was mapped to the node color, for the significantly differentially expressed genes (FDR < 0.05) within the subgraph.

Results

A total of 12,961 genes were available for analysis after filtering out genes with low expression. Of the 1158 nuclear genes listed in the MitoCarta2.0 database, 847 genes (73%) were identified in the Moroccan dataset. The numbers of differentially expressed transcriptome and MitoCarta2.0 database genes, by geographical comparison, are summarized in (Table 1). The number of differentially expressed genes (i.e., the nuclear genes encoding mitochondrial proteins) varied geographically in a pattern that paralleled the gene expression changes of the total transcriptome. The Agadir-Boutroch (urban-rural) and Agadir-Ighrem (urban-rural) comparisons had the greatest number of differentially expressed genes within the MitoCarta2.0 database (76.5% and 81.3%, respectively), while the rural-rural comparison (Boutroch-Ighrem) had the least (10.3%). It is worth noting that the fraction of upregulated genes within the MitoCarta2.0 database exceeded the fraction of upregulated transcriptome genes in both urban-rural comparisons (e.g., 74.0% vs 52.9% in the Agadir-Boutroch comparison).

A more detailed analysis of the Agadir-Boutroch geographical comparison is shown in Fig. 2 and Table 2. In the Agadir-Boutroch urban-rural comparison, 97% of the differentially expressed genes identified in the MitoCarta2.0 database were increased in the urban area (Fig. 2a). The mitochondrial associated protein transcripts (nuclear DNA) with the greatest expression differences are shown in Fig. 2b. The genes with the greatest fold increase change difference were NDUFA4 (logFC 0.89, adj.P.Val 5.77E-14), COX7B (logFC 0.85, adj.P.Val 4.46E-12), and MRPS21 (logFC 0.80, adj.P.Val 1.41E-10). A complete list of differentially expressed MitoCarta2.0 genes for the Agadir-Boutroch comparison is provided in Supplementary Table 1.

Gene Set Enrichment Analysis (GSEA) of the differentially expressed genes within the MitoCarta2.0 database revealed 367 significantly enriched pathways (adjusted p value < 0.05) (Supplementary Table 2), of which oxidative phosphorylation, regulation of insulin secretion, and glucose regulation showed the highest values (Fig. 2c). WebGestalt showed mitochondrial RNA metabolic processing and translational elongation as the most highly enriched biological processes (enrichment ratios 14.0 and 14.8, respectively, FDR < 0.5) (Fig. 2d). SPIA of the differentially expressed MitoCarta2.0 genes revealed four significantly perturbed KEGG disease pathways—three degenerative neurological diseases (Huntington's, Alzheimer's and Parkinson's diseases) and herpes simplex infection (false discover rate corrected P values (PGFdr) < 0.2) (Table 2). A complete list of the SPIA results is provided in Supplementary Table 3. Network visualization of the differentially expressed Mitocarta2.0 genes in the Agadir-Boutroch comparison showed a link between clusters of oxidase-reductase and ribosomal protein biosynthesis genes (Fig. 2e).

Discussion

Urbanization is the most important demographic change during the past century, and is a major divergence from how humans have lived for the past several thousand years (Galea and Vlahov 2005). Although living in a city is usually associated with improved living standards, an untoward effect of urbanization has been an increased prevalence of NCDs, such as cardiovascular diseases, cancer, and respiratory diseases. With no end in sight to the current global epidemic in NCDs, there is a critical need to better understand the pathogenesis of these common conditions, which could lead to new prevention and treatment strategies. In the present study, we

Table 1 Number of differentially expressed genes by geographical comparison

Geographic comparison		Agadir vs. Boutroch (urban-rural)	Agadir vs Ighrem (urban-rural)	Ighrem vs Boutroch (rural-rural)	Average urban rural (Agadir vs average of Boutroch and Ighrem)
Transcriptome	Total genes expressed (%)	12,961 (100)	12,961 (100)	12,961 (100)	12,961 (100)
	No change (%)	5392 (41.6)	4898 (37.8)	11,317 (87.3)	9678 (74.7)
	Upregulated (adjusted p value < 0.05) (%)	6828 (53.9)	7886 (60.8)	134 (1.0)	2321 (17.9)
	Downregulated (adjusted p value < 0.05) (%)	741 (5.7)	177 (1.4)	1510 (11.7)	962 (7.4)
MitoCarta2.0 genes (i.e., nuclear genes encoding mitochondrial proteins)	Total genes expressed (%)	847 (100)	847 (100)	847 (100)	847 (100)
	No change (%)	199 (23.5)	154 (18.2)	760 (89.7)	465 (54.9)
	Upregulated (adjusted p value < 0.05) (%)	627 (74.0)	688 (81.2)	6 (0.7)	340 (40.1)
	Downregulated (adjusted p value < 0.05) (%)	21 (2.5)	5 (0.6)	81 (9.6)	42 (5.0)

examined whether urbanization altered the expression of nuclear genes encoding mitochondrial proteins. Our analysis of gene expression microarray data from individuals living in a rural and urban area of Morocco revealed several interesting findings.

First, it was found that urbanization increases the expression of nuclear genes encoding mitochondrial proteins. This effect was observed in two different urban-rural comparisons and seemed to relate to the degree of environmental difference, as the number of differentially expressed genes in between Ighrem and Boutroch (two rural areas) was only 10.3%. Since Ighrem is mainly of Arab ethnicity, and Boutroch almost exclusively Amazigh, these data indicate that the genetic variations between these two populations only account for a minimal fraction of differentially expressed genes. The majority of changes in nuclear genes encoding mitochondrial protein are thus most likely a result of urbanization. In the urban-rural comparison with the greatest environmental differences, over 75% of the nuclear genes encoding mitochondrial proteins were differentially expressed, with 97% having increased expression in the urban area. As urbanization is associated with changes in multiple environmental stimuli, it is difficult, at this stage, to suggest the factors responsible for these differences in gene expression. Some of the more prominent environmental changes that occur with urbanization include differences in the burden of infectious diseases, diet, crowding, noise, and levels of pollution (Bickler 2000).

Moving forward, it will also be important to determine the mechanism(s) by which urbanization increases the expression of nuclear genes encoding mitochondrial proteins. Based on the existing literature, increased mitochondrial stress in urban areas affecting translational mechanisms, and chromatin-related phenomena that regulates mitochondrial gene expression would seem to be the most likely mechanisms. Mitochondria stress is known to activate nuclear genes via a variety of mechanisms, including OXPHOS dysfunction, defects in mtDNA, and loss of membrane potential (Guha and Avadhani 2013; Arnould et al. 2015). Mitochondrial stress can also activate a mitochondrial-cytosolic response, or an extracellular response mediated by "mitokine" signals (Quiros et al. 2016). The observation that pollution can have profound effects on mitochondrial function supports this hypothesis (Dagher et al. 2006; Gualtieri et al. 2011; Li et al. 2006; Pigneret et al. 2018). Alternatively, the gene expression differences could be caused by a chromatin-related phenomenon—although the two mechanisms are not mutually exclusive. In this scenario, compaction or accessibility of DNA to relevant transcription factors is altered (Andersen and Tost 2018; Chiarella et al. 2020), limiting the expression of the nuclear genes encoding mitochondrial proteins in the rural areas. Given the importance of epigenetic mechanisms in guiding other physiological adaptations, perturbations in chromatin-related processes deserve further study.

An additional important finding was that urbanization potentially affects a broad spectrum of biological processes, with nuclear transcripts encoding mitochondria proteins functioning as an important nexus. In the urban-rural comparison, differentially expressed genes within the MitoCarta2.0 database were associated with more than 350 functionally related gene sets in the Molecular Signatures Database (MSigDB). Similar to the Idaghdour study (Idaghdour et al. 2010), it was found that found urbanization was associated with heavy enrichment of oxidative phosphorylation and ribosomal biosynthesis. Insulin secretion, glucose metabolism, and integration of energy metabolism were other highly enriched pathways. The most heavily enriched biological process were RNA metabolic processes, translational elongation, and mitochondrial transport. Interestingly, our gene interaction

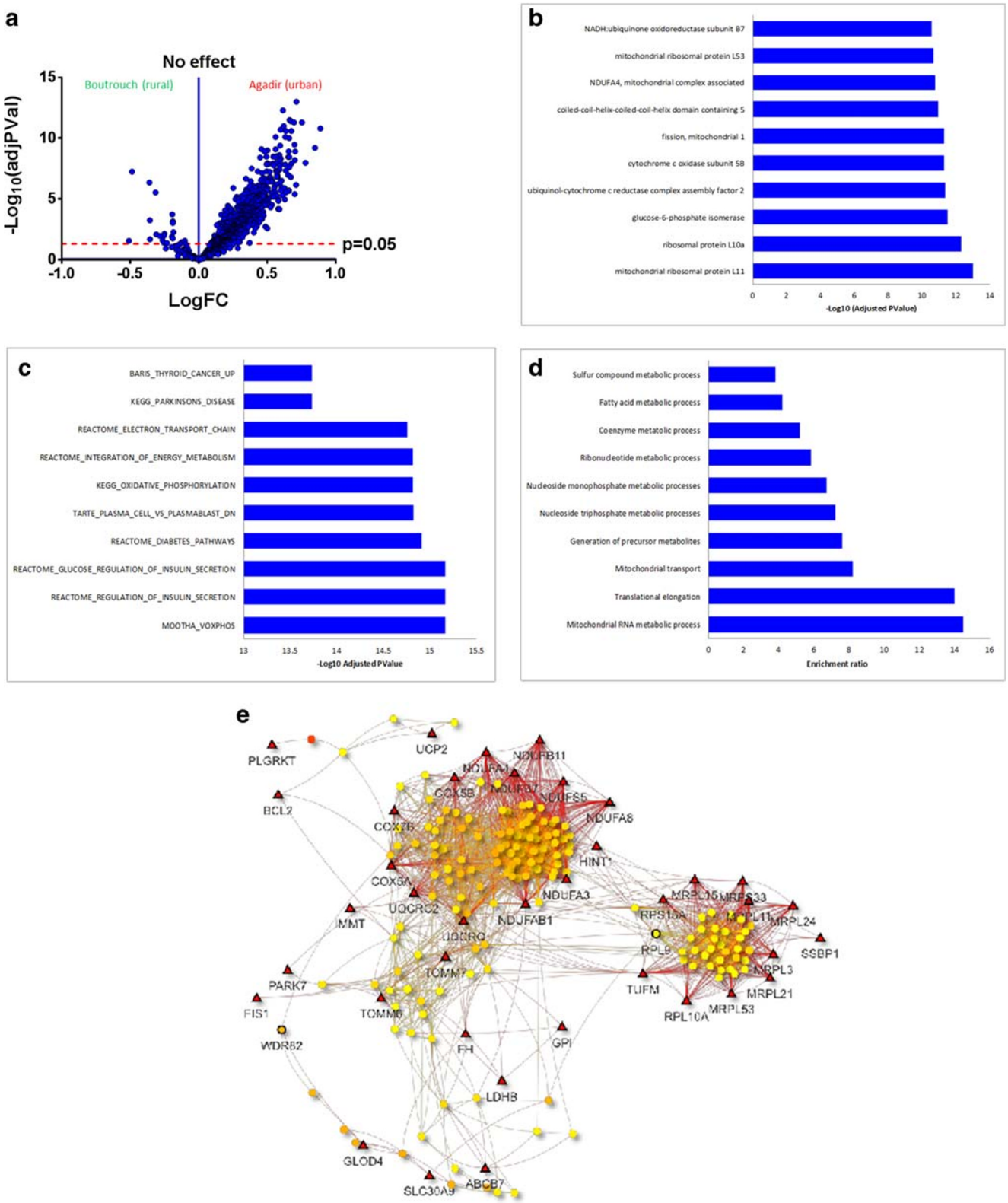

Fig. 2 Analysis of differentially expressed MitoCarta2.0 genes (i.e., nuclear genes encoding mitochondrial proteins) in the Agadir-Boutrouch (urban-rural) geographic comparison. **a** Volcano plot of MitoCarta2.0 genes in Agadir (urban) and Boutrouch (rural). **b** MitoCarta2.0 gene transcripts with the greatest expression differences (top ten). **c** Functionally related gene sets identified by Gene Set Enrichment Analysis (GSEA) (top ten). **d** Biological processes identified by WebGestalt (top ten). **e** Network visualization using Cytoscape and VisJS2jupyter

Table 2 KEGG disease pathways identified by Signaling Pathway Impact Analysis (SPIA)

	Name	ID	pSize	NDE	pNDE	tA	pPERT	pG	pGFdr	pGFWER	Status
1	Huntington's disease	5016	75	70	6.47E-05	− 0.505359048	0.212	0.000167352	0.011714644	0.011714644	Inhibited
2	Alzheimer's disease	5010	70	65	0.000195516	− 0.484697517	0.236	0.000506811	0.01773838	0.03547676	Inhibited
3	Parkinson's disease	5012	73	67	0.000405577	− 1.534221055	0.335	0.001345617	0.0313997721	0.094193164	Inhibited
4	Herpes simplex infection	5168	4	4	0.341836624	− 1.727162773	0.004	0.010384837	0.181734648	0.726938592	Inhibited

psize number of genes in the pathway; *NDE* number of DE genes in pathway; *pNDE* probability to observe at least NDE genes in the pathway; *tA* observed total perturbation accumulation in the pathway; *pPERT* probability to observe a total accumulation more extreme thamn tA by chanve; *pG p* value obtained by combining pNDE and pPERT; *pGFdr* false discovery rate; *pGFWER* Bonferroni adjusted global *p* values; *Status* direction in which the pathway is perturbed in urban population

network analysis revealed a relationship between clusters of oxidative reductase genes and ribosomal protein biosynthesis genes. Given the results to our enrichment analysis, and recent advances in the understanding of mitochondrial gene expression (For review see (Pearce et al. 2017)), we suggest these clusters of genes might represent the mitochondrial ribosome (mitoribosome) and the OXPHOS system. Human mitoribosomes synthesize 13 essential proteins of the oxidative phosphorylation pathway and are composed of between 250 and 300 nuclear encoded proteins (Amunts et al. 2015; Calvo et al. 2016; Pearce et al. 2017; Smith and Robinson 2016). As mitoribosome function is paramount to mitochondrial respiration, and thus critical to cell differentiation, growth, and survival, this could be an important area of future research.

Finally, several disease pathways were associated with urbanization. Specifically, that differentially expressed nuclear genes encoding mitochondrial proteins were associated with inhibition of three degenerative neurological diseases (Huntington's, Alzheimer's and Parkinson's diseases) and herpes simplex virus infection. Because nuclear transcripts encoding mitochondrial proteins are a subset of the genes in the entire pathway, care must be exercised in interpreting whether the disease pathways are either activated or inhibited. As an example, the Alzheimer KEGG disease pathway is inhibited in the urban population when nuclear genes encoding mitochondrial proteins are used in the analysis; but the pathway is activated in the urban population when analyzed using the larger transcriptome (false discover rate corrected *p* values (PGFdr) = 0.007). Perhaps, the most that can be concluded from our Signaling Pathway Impact Analysis (SPIA) is that nuclear transcripts encoding mitochondrial proteins are well represented in several degenerative neurological diseases, and Herpes simplex virus infection pathways. Nevertheless, as mitochondrial dysfunction has been identified in a broad range of disease processes, ranging from neonatal fatalities to adult onset neurodegeneration, and is a likely contributor to cancer and type II diabetes (DiMauro and Schon 2003; Lowell and Shulman 2005; Wallace 2005), this should be an area of fertile research.

Our study did have several limitations, including the incomplete gene expression data in the Moroccan dataset and that the samples were collected from one geographical area (southern Morocco). Repeating the study in other geographic areas and utilizing more comprehensive gene expression tools (e.g., RNAseq) would be the best way to address these limitations. RNAseq analysis has the advantage that it can provide information on gene expression of mitochondrial genes (mtDNA), thus providing important insight into the effect mitochondrial genes might have on nuclear gene expression. Knowing the expression of mitochondrial genes (i.e., the subunits of the OXPHOS system, 22 transfer RNAs, and two ribosomal RNAs) in different environments could help answer this question. Another limitation of our study is that it was based solely on the gene expression in peripheral white blood cells. Although this may not be a problem, as there is growing evidence that the blood transcriptome dynamically reflects system wide biology (Liew et al. 2006).

In conclusion, our study suggests a link between urbanization in Morocco and changes in the expression of nuclear genes encoding mitochondrial proteins. In doing so, it should bring to bear extensive literature on the role of mitochondria in cellular function, and thus be helpful in better understanding how urbanization impacts mitochondrial function, and ultimately the role it might have in the origins of NCDs. Recent themes in mitochondrial research that might be relevant to the urban-rural transition include the relationship between mitochondria and the innate immune system (Garaude 2018; Sander and Garaude 2018), and the emerging concept that a network of organelles helps maintain cellular homeostasis (Gilkerson 2018). In the latter, mitochondrial dynamics are thought to be intrinsically linked to bioenergetics and adenosine triphosphate (ATP) production, and mediate cell-wide signaling networks (Gilkerson 2018). The urban-rural paradigm offers an important opportunity to investigate these mechanisms in humans, and how they might relate to the rising rates of NCDs worldwide.

Funding information This study was supported by the Fogarty International Center, National Institutes of Health (NIH), grant number R25TW011216 and the University of California San Diego (UCSD), Clinical and Translational Research Institute (CTRI) grant number UL1TR001442. The authors thank Barbara Rho for her excellent editing of the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interests.

References

Amegah AK, Agyei-Mensah S (2017) Urban air pollution in Sub-Saharan Africa: time for action. Environ Pollut 220(Pt A):738–743. https://doi.org/10.1016/j.envpol.2016.09.042

Amunts A, Brown A, Toots J, Scheres SHW, Ramakrishnan V (2015) Ribosome. The structure of the human mitochondrial ribosome. Science 348:95–98. https://doi.org/10.1126/science.aaa1193

Andersen GB, Tost J (2018) A summary of the biological processes, disease-associated changes, and clinical applications of DNA methylation. Methods Mol Biol 1708:3–30. https://doi.org/10.1007/978-1-4939-7481-8_1

Arnould T, Michel S, Renard, P (2015) Mitochondria Retrograde Signaling and the UPR mt: Where Are We in Mammals? Int J Mol Sci 16(8):18224–18251. https://doi.org/10.3390/ijms160818224

Ayeni FA, Biagi E, Rampelli S, Fiori J, Soverini M, Audu HJ, Turroni S (2018) Infant and adult gut microbiome and metabolome in rural Bassa and urban settlers from Nigeria. Cell Rep 23:3056–3067. https://doi.org/10.1016/j.celrep.2018.05.018

Bickler SW (2000) Non-communicable diseases: is their emergence in industrialized societies related to changes in neuroendocrine function? Med Hypotheses 54:825–828. https://doi.org/10.1054/mehy.1999.0959

Bickler SW, Lizardo RE, De Maio A (2015) The transition from a rural to an urban environment alters expression of the human Ebola virus receptor Neiman-pick C1: implications for the current epidemic in West Africa. Cell Stress Chaperones 20:203–206. https://doi.org/10.1007/s12192-014-0557-9

Bickler SW, Lizardo E, Cauvi DM, De Maio A (2016) The transition from a rural to an urban environment in Africa alters G protein-coupled receptor signaling. Med Hypotheses 95:49–53. https://doi.org/10.1016/j.mehy.2016.08.004

Bickler SW, Wang A, Amin S, Halbach J, Lizardo R, Cauvi DM, De Maio A (2018) Urbanization in sub-Saharan Africa: declining rates of chronic and recurrent infection and their possible role in the origins of non-communicable diseases. World J Surg 42:1617–1628. https://doi.org/10.1007/s00268-017-4389-5

Buchanan FG, DuBois RN (2006) Emerging roles of beta-arrestins. Cell Cycle 5:2060–2063. https://doi.org/10.4161/cc.5.18.3212

Caldwell JC (2001) Population health in transition. Bull World Health Organ 79:159–160 retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11242823

Calvo SE, Clauser KR, Mootha VK (2016) MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. Nucleic Acids Res 44:D1251–D1257. https://doi.org/10.1093/nar/gkv1003

Chen J, Bardes EE, Aronow BJ, Jegga AG (2009) ToppGene suite for gene list enrichment analysis and candidate gene prioritization. Nucleic Acids Res 37:W305–W311. https://doi.org/10.1093/nar/gkp427

Chen M, Zhang H, Liu W, Zhang W (2014) The global pattern of urbanization and economic growth: evidence from the last three decades. PLoS One 9:e103799. https://doi.org/10.1371/journal.pone.0103799

Chiarella AM, Lu D, Hathaway NA (2020) Epigenetic control of a local chromatin landscape. Int J Mol Sci 21:943. https://doi.org/10.3390/ijms21030943

Coker E, Kizito S (2018) A narrative review on the human health effects of ambient air pollution in Sub-Saharan Africa: an urgent need for health effects studies. Int J Environ Res Public Health:15. https://doi.org/10.3390/ijerph15030427

Cowen L, Ideker T., Raphael BJ, Sharan R (2017) Network propagation: a universal amplifier of genetic associations. Nat Rev Genet 18(9): 551-562. https://doi.org/10.1038/nrg.2017.38

Dagher Z, Garcon G, Billet S, Gosset P, Ledoux F, Courcot D, Aboukais A, Shirali P (2006) Activation of different pathways of apoptosis by air pollution particulate matter (PM2.5) in human epithelial lung cells (L132) in culture. Toxicology 225:12–24. https://doi.org/10.1016/j.tox.2006.04.038

Davis S, Meltzer PS (2007) GEOquery: a bridge between the gene expression omnibus (GEO) and BioConductor. Bioinformatics 23: 1846–1847. https://doi.org/10.1093/bioinformatics/btm254

DiMauro S, Schon EA (2003) Mitochondrial respiratory-chain diseases. N Engl J Med 348:2656–2668. https://doi.org/10.1056/NEJMra022567

Dunning MLA, Eldridge M (2015) illuminaHumanv3.db: Illumina HumanHT12v3 annotation data (chip illuminaHumanv3). R package version 1.26.0. Retrieved from http://bioconductor.org/.

Eckert S, Kohler S (2014) Urbanization and health in developing countries: a systematic review. World Health Popul 15:7–20 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/24702762

Galea S, Vlahov D (2005) Urban health: evidence, challenges, and directions. Annu Rev Public Health 26:341–365. https://doi.org/10.1146/annurev.publhealth.26.021304.144708

Garaude J (2018) Reprogramming of mitochondrial metabolism by innate immunity. Curr Opin Immunol 56:17–23. https://doi.org/10.1016/j.coi.2018.09.010

Gentleman R, Huber W, Hahne F (2018) genefilter: methods for filtering genes from high-throughput experiments. R package version 1.64.0. Retrieved from http://bioconductor.org/. https://doi.org/10.18129/B9.bioc.genefilter

Gilkerson R (2018) A disturbance in the force: cellular stress sensing by the mitochondrial network. Antioxidants (Basel) 7. https://doi.org/10.3390/antiox7100126

Gong P, Liang S, Carlton EJ, Jiang Q, Wu J, Wang L, Remais JV (2012) Urbanisation and health in China. Lancet 379:843–852. https://doi.org/10.1016/S0140-6736(11)61878-3

Gualtieri M, Ovrevik J, Mollerup S, Asare N, Longhin E, Dahlman HJ, Camatini M, Holme JA (2011) Airborne urban particles (Milan winter-PM2.5) cause mitotic arrest and cell death: effects on DNA, mitochondria, AhR binding and spindle organization. Mutat Res 713:18–31. https://doi.org/10.1016/j.mrfmmm.2011.05.011

Guha M, Avadhani NG. (2013) Mitochondrial retrograde signaling at the crossroads of tumor bioenergetics, genetics and epigenetics. Mitochondrion 13(6):577–591. https://doi.org/10.1016/j.mito.2013.08.007

Hanzelmann S, Castelo R, Guinney J (2013) GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 14: 7. https://doi.org/10.1186/1471-2105-14-7

Harpham T (2009) Urban health in developing countries: what do we know and where do we go? Health Place 15:107–116. https://doi.org/10.1016/j.healthplace.2008.03.004

Htet AS, Kjollesdal MK et al (2017) Lipid profiles and determinants of total cholesterol and hypercholesterolaemia among 25-74 year-old urban and rural citizens of the Yangon Region, Myanmar: a cross-sectional study. BMJ Open 7:e017465. https://doi.org/10.1136/bmjopen-2017-017465

Idaghdour Y, Storey JD, Jadallah SJ, Gibson G (2008) A genome-wide gene expression signature of environmental geography in leukocytes of Moroccan Amazighs. PLoS Genet 4:e1000052. https://doi.org/10.1371/journal.pgen.1000052

Idaghdour Y, Czika W, Shianna KV, Lee SH, Visscher PM, Martin HC, Miclaus K, Jadallah SJ, Goldstein DB, Wolfinger RD, Gibson G (2010) Geographical genomics of human leukocyte gene expression variation in southern Morocco. Nat Genet 42:62–67. https://doi.org/10.1038/ng.495

Jamison DT, Mosley WH (1991) Disease control priorities in developing countries: health policy responses to epidemiological change. Am J Public Health 81(1):15–22. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/1983911

Jha AR, Davenport ER, Gautam Y, Bhandari D, Tandukar S, Ng KM, Fragiadakis GK, Holmes S, Gautam GP, Leach J, Sherchand JB, Bustamante CD, Sonnenburg JL (2018) Gut microbiome transition across a lifestyle gradient in Himalaya. PLoS Biol 16:e2005396. https://doi.org/10.1371/journal.pbio.2005396

Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res 45:D353–D361. https://doi.org/10.1093/nar/gkw1092

Kavishe B, Vanobberghen F, Katende D, Kapiga S, Munderi P, Baisley K, Biraro S, Mosha N, Mutungi G, Mghamba J, Hughes P, Smeeth L, Grosskurth H, Peck R (2019) Dyslipidemias and cardiovascular risk scores in urban and rural populations in north-western Tanzania and southern Uganda. PLoS One 14:e0223189. https://doi.org/10.1371/journal.pone.0223189

Lane N (2005) Power, sex, suicide: mitochondria and the meaning of life: Oxford University Press

Li Z, Hyseni X, Carter JD, Soukup JM, Dailey LA, Huang YC (2006) Pollutant particles enhanced H2O2 production from NAD(P)H oxidase and mitochondria in human pulmonary artery endothelial cells. Am J Physiol Cell Physiol 291:C357–C365. https://doi.org/10.1152/ajpcell.00365.2005

Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P (2015) The molecular signatures database (MSigDB) hallmark gene set collection. Cell Syst 1:417–425. https://doi.org/10.1016/j.cels.2015.12.004

Liew CC, Ma J, Tang HC, Zheng R, Dempsey AA (2006) The peripheral blood transcriptome dynamically reflects system wide biology: a potential diagnostic tool. J Lab Clin Med 147:126–132. https://doi.org/10.1016/j.lab.2005.10.005

Lokmer A, Aflalo S, Amougou N, Lafosse S, Froment A, Tabe FE, Poyet M, Groussin M, Said-Mohamed R, Ségurel L (2020) Response of the human gut and saliva microbiome to urbanization in Cameroon. Sci Rep 10:2856. https://doi.org/10.1038/s41598-020-59849-9

Lowell BB, Shulman GI (2005) Mitochondrial dysfunction and type 2 diabetes. Science 307:384–387. https://doi.org/10.1126/science.1104343

Monlun M, Hyernard C, Blanco P, Lartigue L, Faustin B (2017) Mitochondria as molecular platforms integrating multiple innate immune signalings. J Mol Biol 429:1–13. https://doi.org/10.1016/j.jmb.2016.10.028

Mortality, GBD, Causes of Death C (2016) Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 388(10053):1459–1544. https://doi.org/10.1016/S0140-6736(16)31012-1

Murray CJ, Lopez AD (1997) Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. Lancet 349(9063):1436–1442. https://doi.org/10.1016/S0140-6736(96)07495-8

Naquet P, Giessner C, Galland F (2016) Metabolic adaptation of tissues to stress releases metabolites influencing innate immunity. Curr Opin Immunol 38:30–38. https://doi.org/10.1016/j.coi.2015.10.005

Nath AP, Arafat D, Gibson G (2012) Using blood informative transcripts in geographical genomics: impact of lifestyle on gene expression in Fijians. Front Genet 3:243. https://doi.org/10.3389/fgene.2012.00243

Pearce SF, Rebelo-Guiomar P, D'Souza AR, Powell CA, Van Haute L, Minczuk M (2017) Regulation of mammalian mitochondrial gene expression: recent advances. Trends Biochem Sci 42:625–639. https://doi.org/10.1016/j.tibs.2017.02.003

Pigneret M, Roussel D, Hervant F (2018) Anaerobic end-products and mitochondrial parameters as physiological biomarkers to assess the impact of urban pollutants on a key bioturbator. Environ Sci Pollut Res Int 25:27225–27234. https://doi.org/10.1007/s11356-018-2756-x

Quiros PM., Mottis A, Auwerx J (2016) Mitonuclear communication in homeostasis and stress. Nat Rev Mol Cell Biol 17(4):213–226. https://doi.org/10.1038/nrm.2016.23

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015) Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43:e47. https://doi.org/10.1093/nar/gkv007

Rosenthal SB, Len J, Webster M, Gary A, Birmingham A, Fisch KM (2018) Interactive network visualization in Jupyter notebooks: visJS2jupyter. Bioinformatics 34:126–128. https://doi.org/10.1093/bioinformatics/btx581

Sander LE, Garaude J (2018) The mitochondrial respiratory chain: a metabolic rheostat of innate immune cell-mediated antibacterial responses. Mitochondrion 41:28–36. https://doi.org/10.1016/j.mito.2017.10.008

Satterthwaite D, McGranahan G, Tacoli C (2010) Urbanization and its implications for food and farming. Philos Trans R Soc Lond Ser B Biol Sci 365:2809–2820. https://doi.org/10.1098/rstb.2010.0136

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504. https://doi.org/10.1101/gr.1239303

Smith AC, Robinson AJ (2016) MitoMiner v3.1, an update on the mitochondrial proteomics database. Nucleic Acids Res 44:D1258–D1261. https://doi.org/10.1093/nar/gkv1001

Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ, von Mering C (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 43(Database issue):D447–D452. https://doi.org/10.1093/nar/gku1003

Tarca AL KP, Draghici S (2018) SPIA: Signaling Pathway Impact Analysis (SPIA) using combined evidence of pathway over-representation and unusual signaling perturbations. R package version 2.34.0. Retrieved from https://www.bioconductor.org/

United Nations (2019) World Urbanization Prospects: The 2018 Revision. Retrieved from https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf

Vasunilashorn S, Crimmins EM, Kim JK, Winking J, Gurven M, Kaplan H, Finch CE (2010) Blood lipids, infection, and inflammatory markers in the Tsimane of Bolivia. Am J Hum Biol 22:731–740. https://doi.org/10.1002/ajhb.21074

Wallace DC (2005) Mitochondria and cancer: Warburg addressed. Cold Spring Harb Symp Quant Biol 70:363–374. https://doi.org/10.1101/sqb.2005.70.035

Wang J, Vasaikar S, Shi Z, Greer M, Zhang B (2017) WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. Nucleic Acids Res 45:W130–W137. https://doi.org/10.1093/nar/gkx356

Affiliations

Stephen W. Bickler[1,2,3] · James M. Prieto[4] · David M. Cauvi[2,3] · Victor De Cos[1] · Chanond Nasamran[5] · Emmanuel Ameh[6] · Said Amin[7] · Sneha Nicholson[1] · Hena Din[1] · Ana Olga Mocumbi[8,9] · Emilia Virginia Noormahomed[9] · Guillermo Tellez-Isaias[10] · Kathleen M. Fisch[5] · Antonio De Maio[2,3,11]

1 Division of Pediatric Surgery, Rady Children's Hospital—University of California San Diego, 3030 Children's Way, San Diego, CA 92123, USA

2 Department of Surgery, School of Medicine, University of California San Diego, La Jolla, CA 92093, USA

3 Center for Investigations of Health and Education Disparities, University of California San Diego, La Jolla, CA 92093, USA

4 Department of Surgery, Naval Medical Center San Diego, San Diego, CA, USA

5 Center for Computational Biology and Bioinformatics, University of California San Diego, La Jolla, CA 92093, USA

6 Department of Pediatric Surgery, National Hospital, Abuja, Nigeria

7 Department of Histopathology, National Hospital, Abuja, Nigeria

8 Instituto Nacional de Saúde, Maputo, Mozambique

9 Department of Microbiology, Faculty of Medicine, Universidade Eduardo Mondlane (UEM), Maputo, Mozambique

10 Department of Poultry Science, University of Arkansas, Fayetteville, AR, USA

11 Department of Neurosciences, School of Medicine, University of California San Diego, La Jolla, CA 92093, USA