Ancient urbanization, historical prevalence of pathogens and the ADH1B*48His allele frequency

Shcherbak Andrey[1] (LCSR HSE), Anna A. Kim (VGG RAS)[2], Borinskaya Svetlana[3] (VIGG RAS)

This paper aims to make a contribution to a broad gene-culture coevolution theoretical framework. Urbanization is often seen as a proxy for complex social organization: the size of the largest city in a region has been a function of the scale of political organization (Morris, 2013: 165).. Urban settlements with population more than 1000 residents can be dated as early as 7000 BCE (Morris, 2013). Increased population density may drive infectious disease rate upward. Therefore, ancient urbanization may be associated with history of pathogens. In this paper we studied the potential impact of ancient urbanization on selection of genes responsible for resistance to infectious diseases, using the ADH1B*48His as an example.

This gene is encoding an enzyme alcohol dehydrogenase that oxidizes ethanol to acetaldehyde. The gene is polymorphic. A gene variant called ADH1B*48His is reported to be associated with faster accumulation of acetaldehyde in the blood. That results "flushing syndrome" – elevated blood flow, dizziness, accelerated heart rate, sweating and nausea (Mulligan et al., 2003); and due to these symptoms carriers of the allele are reported to consume less alcohol and to be less exposed to alcohol abuse and alcohol dependence. There is a substantial variation among populations on the ADH1B*48His allele frequency (Borinskaya et al., 2009) and strong evidence of selection of this gene variant at least in some populations (Li et al., 2011).

There is a vast literature on the relationship between ancient urbanization and the rise and spread of infections: ancient urbanization is strongly associated with the increase in pathogen load. Urban settlements show larger populations and much higher population density what facilitates emergence and spread of particular infections, or 'crowd diseases' (Clark, 2010; Barnes, 2005; Cook, 2013; Shug et al., 2013; Miksic, 1999).

[1] Senior Research Fellow, Laboratory for Comparative Social Research, National Research University Higher School of Economics. 47a, Rimskogo-Korsakova pr., St.-Petersburg, Russia 190068. E-mail: ascherbak@hse.ru

[2] Junior Research Fellow, The Vavilov Institute of General Genetics, Russian Academy of Sciences. 3, Gubkina str., Moscow, Russia, 119991. E-mail: Anna.Kim@vigg.ru

[3] Senior Research Fellow, The Vavilov Institute of General Genetics, Russian Academy of Sciences. 3, Gubkina str., Moscow, Russia, 119991. E-mail: borinskaya@vigg.ru

It can be thus defined as a reciprocal effect: on the one hand, ancient urbanization is likely to lead to explosion of infectious diseases that became the main mortality reason for population.. On the other hand, earlier urbanization means early start of selection of genes associated with natural resistance to these infections.

Infectious diseases affect natural selection: carriers of genes resistant to diseases are 'winners' of evolution. Therefore, we guess that ancient urbanization, followed by the increase in infectious diseases may cause higher frequencies of gene variants that are associated with resistance to infectious diseases.

We found at least one paper that supports our argument. Ancient urbanization is reported to be connected with the frequency of an allele (SLC11A1 1729 + 55del4) associated with natural resistance to intracellular pathogens such as tuberculosis and leprosy (Barnes et al., 2010: 2010). They found a highly significantly correlation with duration of urban settlement — populations with a long history of living in towns are better adapted to resisting these infections. We suggest a similar causal mechanism in our case.

Data

Data on **ancient urbanization** and index of historical pathogens prevalence suggested by Murray and Schaller (2010). Using data from the OpenHistory Project (OpenHistory Project) and Chandler (1987) we construct two variables: a) *Number of cities at 650 BCE* – the number of cities that existed in a particular geopolitical area at 650 BCE (not including those one that were destroyed or abandoned before this date); b) *Urban population estimate at 650 BCE* – an approximate estimate of urban population in all existing cities in a particular geopolitical area. Then we created another variable - *Urban history* – measured as '2000 - year of the first urban settlement' (for 40 populations/ regions).

To measure **historical prevalence of pathogens**, we use the both 7 and 9 items indices suggested by Murray and Schaller (2010). It is *pathogens prevalence 7 and 9* variables. Murray and Schaller (2010) composed their historical pathogen prevalence indices based on two sources – Rodenwaldt's and Bader's *World-atlas of epidemic diseases* (Rodenwald and Bader, 1952, 1961) and Simmons et al.'s *Global epidemiology*. (Simmons et al., 1944). We referred to these editions to obtain data on each infections: leishmaniasis, schistosomiasis, leprosy, malaria, typhus, filariasis, dengue, and tuberculosis. A 4-point coding scheme was employed: 0 = completely absent or never reported, 1 = rarely reported, 2 = sporadically or moderately reported, 3 = present at severe levels or epidemic levels at least once. Based on this data we created

variables on level of prevalence of these infections following the same coding procedure as in the Murray and Schaller's paper (2010) for required populations.

Data on the **ADH1B*48His allele frequency** were taken from the ALFRED database. Cases were selected from the ALFRED database by: a) Presence of at least 50 individuals in the sample; b) Ability to match this ethnic group with an existing polity; c) exclusion of 'mixed' populations (like 'Indians (mixed)', 'South Africans' and 'Taiwanese'). Finally, some cases were omitted because data on pathogens prevalence and urbanization are unavailable. No population from the New World was included. Forty populations meet these requirements. These populations are: Moroccans, Saharawi, Iranian, Kazakh, Kyrgyz, Mongolians, Tajik, Uzbek, Han Chinese, Japanese, Koreans, Filipino, Thai, Malaysians, Papuan New Guinea, Turks, Belorussian, Croatian, Czech, Danes, English, Finns, French, German, Greeks, Hungarian, Irish, Italians, Norwegians, Poles, Romanians, Russians, Slovaks, Spaniards, Swedes, Ukrainian, Armenian, Turkman, Cambodians, and Vietnamese.

To test that correlations are not the result of interregional differences (Europe vs. Asia) we also study interregional correlations. To make such a robustness test, we used 61 Central, East and South-East Asian populations with the highest frequencies of the ADH1B*48His allele (values higher than 0.250). It is mostly Chinese regional/provincial level (48 populations); we also add populations from Central Asia (Tajiks and Uzbeks), East Asia (Koreans, Japanese, Laoloum, Vietnamese, Thai, Mongolians, Malaysians, and Filipino) and from the Pacific (Australian Aborigines and Maori) to this sample. We take these data from Peng et al. (2010) and the ALFRED database

Methods

We use a four-step correlation test to find association between ancient urbanization, historical prevalence of pathogens and the ADH1B*48His allele frequency.

Results.

Firstly, we correlated (Pearson's correlation) index of historical prevalence of pathogens (both 7 and 9 items) with *Number of cities at 650 BCE* and *Urban population estimate at 650 BCE*, on 17 geopolitical regions. Urban population correlates with the 9-item Index of historical diseases ($r=0.583$, $p=0.013$), and it's almost significant for the 7-items index ($r=0.480$, $p=0.051$). We also found significant correlations between number of ancient cities and both 9-item ($r=0.594$, $p=0.012$) and 7-item indices ($r=0.563$, $p=0.018$) of historical diseases. On the level of separate infections in this sample, both *Urban estimate* and *N of cities* are both correlated with leishmaniasis ($r=0.608$, $p=0.010$; $r=0,698$, $p<0.01$ respectively), dengue ($r=0.771$, $p<0.01$;

r=0.800, p=<0.01 respectively), and schistosomiasis (r=0.705, p<0.01; r=0.593, p=0,012 respectively); *N of cities* is marginally correlated with leprosy (r=0.524, p=0.045) and typhus (r=0.496, p=0.043).

Secondly, we correlated *Urban history* with the *Index of historical prevalence of pathogens* (Pearson's correlation), and then separate infections (Spearmen's correlation), on 40 populations. We find strong and negative correlation between *urban history* and the *index of historical prevalence of pathogens* (r= 0.481, p<0.01).

The connection between urbanization and infections is even stronger if one focuses on the separate infections' level. The relationship between Urban history and leishmaniasis (r = 0.658, p<0.01) is extremely strong and powerful. These findings are consistent with previous results that reveal the relationship between urbanization and distribution of leishmaniasis in the modern world, especially in developing countries (e.g., WHO leishmaniasis, 2014; Desjeux, 2001). These findings support our argument on the significant relationship between urbanization history and pathogens' load.

Thirdly, we correlated *the ADH1B*48His allele frequency* with the *index of historical prevalence of pathogens* (Pearson's correlation). The correlation index is positive and significant (r=0.605, p<0.01). When we test separate infections (Spearmen's correlation), we find stronger correlations. The most significant infections are dengue (r=0.696, p=<0.01), malaria (r=0.643, p<0.01) and filariasis (r =0.543, p<0.01).

Other correlations are also significant: leprosy (r=0.581, p<0.01, N=34), leischmanias (r = 0.516, p<0.01), and schistosomiasis (r = 0.478, p<0.01). Then we correlated *the ADH1B*48His allele frequency* with *history of urbanization*. Again, we found positive significant correlation (r =0.354, p = 0.025). In our opinion, it provides an evidence for our main argument on relationship between urbanization, historical pathogen prevalence and some allele frequencies.

Finally, we limited our scope only by East Asian populations. For these populations there is some evidence of ADH1B*48His allele selection (Li et al., 2011). To do that we pooled together data on the ADH1B*48His frequency in East Asian populations, including regional/provincial level in China (ALFRED; Peng et al., 2010): 61 populations, including 48 regional/provincial Chinese populations. We correlated of the ADH1B*48His allele frequency with distribution of separate infections (Spearmen's correlation) (leischmaniasis, schistosomiasis, malaria, typhus, filariasis, dengue and ascariasis). Our results support previous findings; almost all correlations are strong and significant: filariasis (r = 0.653, p<0.01), ascariasis (r = 0.495, p<0.01), dengue (r = 0.487, p<0.01), malaria (r = 0.472, p<0.01), schistosomatoiasis (r = 0.467, p<0.01) and leishmaniasis (r = 0.305, p=0.017).

Conclusions. We find strong and significant correlations between ancient urbanization, historical prevalence of pathogens and ADH1B*48His allele frequency. Our results согласуются с findings of Fumagalli et al (2011), who found correlation between frequency of ADH1B*48His with pathogen load, and are complimentary to the findings of Peng et al. (2010). This study revealed the close association of the ADH1B*48His allele frequency and history of rice domestication among East Asian populations. The authors suggested that this association could be interpreted as a selective pressure against increased alcohol intake которое возникло как by-product of rice domestication. We provide an alternative explanation: the very process of irrigation-based agriculture where peasants had to spend many hours with naked feet in slack water could lead to an increased exposure of these peasants to various infectious diseases. The exposure to infectious diseases, not increased consumption of alcohol beverages would create a selective pressure on populations.

We hope that this paper can make at least two contributions to the field. Firstly, we show that variation in the ADH1B*48His allele frequency among populations can be at least partially explained by trends in ancient urbanization, the followed increase in pathogen load and selection pressure for resistance to infectious diseases. Initially we use the index of historical pathogens prevalence and then we examine the effect of separate infections it consists of. Secondly, on the theoretical level we suggest a conceptual framework of allele frequency variation among populations. We argue that this variation may be explained by social factors like emergence of the first urban settlements and states. Urbanization means the rise in population density and radical increase in infectious diseases that could become a powerful mechanism of genetic selection and genetic mutations.

References

ALFRED – The Allele Frequency Database http://alfred.med.yale.edu/

Barnes E., 2005. *Diseases and Human Evolution*. University of New Mexico Press, Albuquerque, Mexico.

Barnes I. et al., 2010. Ancient Urbanization Predicts Genetic Resistance to Tuberculosis. // *Evolution*, Vol. 65, #3. Pp. 842–848.

Borinskaya et al., 2009. Distribution of the Alcohol Dehydrogenase ADH1B*47His Allele in Eurasia. // *The American Journal of Human Genetics*. Vol. 84, №1. Pp. 89–92.

Chandler T., 1987. *Four Thousand Years of Urban Growth. An Historic Census*. Lewston, NY. Edwin Mellen Press.

Clark, D.P. 2010. *Germs, Genes, and Civilization: How Epidemics Shaped Who We Are Today*. Financial Times Press/Pearson.

Cook J., 2013. The Natural Selection of Infectious Disease Resistance and Its Effect on Contemporary Health. Working Paper. Available at: https://dl.dropboxusercontent.com/u/96311558/hla_het_6-3-13.pdf

Desjeux P., 2001**.** The increase in risk factors for leishmaniasis worldwide**. //** *Transactions of the Royal Society of Tropical Medicine and Hygiene*. Vol. 95 (3). Pp. 239–243.

Fumagalli M, Sironi M, Pozzoli U, Ferrer-Admetlla A, Pattini L, Nielsen R., 2011. Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLoS Genet*. 7(11):e1002355

Li D., Zhao H., Gelernter J., 2011. Strong Association of the Alcohol Dehydrogenase 1B Gene (ADH1B) with Alcohol Dependence and Alcohol-Induced Medical Diseases. // *Biological Psychiatry*, 2011. Vol.70. Pp. 504–512.

Miksic J., 1999. Water, Urbanization, and Disease in Early Indonesia, in Bacus, Elizabeth A. and Lucero, Lisa J. (eds.) *Complex Polities in the Ancient Tropical World*. *Archaeological Papers of the American Anthropological Association*, No. 9. Arlington, VA: American Anthropological Association. Pp. 167-184.

Morris I., 2013. *The Measure of Civilization: How Social Development Decides the Fate of Nations*. Princeton: Princeton University Press.

Mulligan C. et al., 2003. Allelic variation at alcohol metabolism genes (ADH1B, ADH1C, ALDH2) and alcohol dependence in an American Indian population. // *Human genetics*, Vol. 113, Issue 4. Pp 325-336.

Murray D., Schaller M., 2010. Historical Prevalence of Infectious Deceases within 230 Geopolitical Regions: a Tool for Investigating Origins of Culture. // *Journal of Cross-Cultural Psychology*, Vol. 41, №1.Pp. 99 –108.

OpenHistory Project - http://openhistory.net/

Peng Yi, Hong Shi, Xue-bin Qi, Chun-jie Xiao, Hua Zhong, Run-lin Z Ma, Bing Su (2010). The ADH1B Arg47His polymorphism in East Asian populations and expansion of rice domestication in history. // *BMC Evolutionary Biology*, 10:15.

Rodenwaldt, E., & Bader, R. E., 1952-1961. *World-atlas of epidemic diseases*. Hamburg, Germany: Falk-Verlag.

Simmons, J. S., Whayne, T. F., Anderson, G. W., & Horack, H. M., 1944. *Global epidemiology*. Philadelphia: J. B. Lippincott.

Schug R., Blevins KE, Cox B, et al., 2013. Infection, Disease, and Biosocial Processes at the End of the Indus Civilization. // *PLoS ONE* 8(12).

WHO Leishmaniasis, 2014. World Health Organization Leishmaniasis Fact Sheet N 375. January 2014. Available at: http://www.who.int/mediacentre/factsheets/fs375/en/ (accessed May 30 2014)