

**Univerzita Karlova, Přírodovědecká fakulta
Katedra antropologie a genetiky člověka**

**Charles University, Faculty of Science
Department of Anthropology and Human Genetics**

Doktorský studijní program: Antropologie a genetika člověka
Ph. D. study program: Anthropology and human genetics

Autoreferát disertační práce
Summary of the Ph. D. thesis



Koevoluční aspekty genetické diverzity:
mléčná produkce a laktázová persistence

Co-evolutionary aspects of genetic diversity:
milk production and lactase persistence

Mgr. Edita Priehodová

Školitel/Supervisor: doc. Mgr. Viktor Černý, Dr.

Praha, 2017

Abstrakt

Laktázovou persistencí (LP) nazýváme geneticky podmíněnou schopnost trávit laktózu (mléčný cukr) v dospělosti. Laktóza je štěpena enzymem tenkého střeva – laktázou, jehož produkce u savců po odstavení ustává. V případě LP je laktáza produkována po celý život, čímž je umožněno bez obtíží přijímat větší množství čerstvého mléka. LP je doložena jen u některých lidských populací. Celosvětově je její rozšíření nerovnoměrné, nejčastější je v severní Evropě a u pastevců aridních oblastí Afriky a Arábie. LP souvisí se vznikem zemědělství před 10 000 lety a počátky mléčné produkce. V určitých populacích nacházíme signál silné pozitivní selekce pro LP, zřejmě podmíněnou různými faktory v závislosti na různých životních podmínkách. LP způsobují jednobodové polymorfismy nacházející se zhruba 14 kb *upstream* od laktázového genu. Dnes je známo pět různých pro-LP variant, které vznikly a rozšířily se zcela nezávisle.

Jedním z možných center vzniku LP je Arabský poloostrov, kde se poprvé objevila a rozšířila pro-LP varianta –13915*G. V této práci bylo studováno rozšíření pro-LP mutací u populací jižní Arábie a afrických Arabů. U všech analyzovaných souborů byl potvrzen výskyt –13915*G. V jižní Arábii pak u většiny souborů ve frekvenci vyšší než 50 %. U afrických Arabů patřících k pastevcům velbloudů byl zaznamenán vyšší výskyt této mutace (44,4–76 %) oproti africkým Arabům, jež přešli k chovu hovězího dobytka či usedlému zemědělství (1,4–25,5 %). Výsledky podporují hypotézu předchozích studií, že –13915*G vznikla a prvně se rozšířila v Arábii pravděpodobně v souvislosti s domestikací velblouda. Výsledky odhadu stáří a expanze pak ukazují, že –13915*G se z Arábie do Afriky rozšířila migracemi probíhajícími od 7. století našeho letopočtu.

Abstract

Lactase persistence (LP) is genetically determined ability to digest lactose (milk sugar) in adulthood. Lactose is digested by the intestinal enzyme – lactase. In mammals, production of lactase declines after the weaning period. In case of the LP, the lactase is produced during whole life and thus it is possible to drink larger amounts of fresh milk without digestive problems. However, LP is documented only in some human populations. Its worldwide distribution is unequal, the highest rates are found in northern Europe and in pastoralists living in arid regions of Africa and Arabia. The origin of this trait is related to the emergence of agriculture 10,000 years ago, and the beginnings of milk production. In some populations, there is a strong signal of positive selection for the LP caused by various factors depending on various living conditions. LP is associated with single nucleotide polymorphisms (SNPs) located approximately 14 kb upstream from lactase gene. Nowadays we know five different SNPs which cause the LP and they have probably different origin and have expanded independently.

One of the possible centres of the LP origin is the Arabian Peninsula, where the LP variant –13,915*G probably first appeared and spread. In this thesis, occurrences of different LP mutations were investigated in the populations of southern Arabia and African Arabs. In all analysed datasets, the presence of the –13,915*G has been confirmed. The frequency of the –13,915*G in southern Arabia was higher than 50% in most of analysed datasets. Higher incidence of this mutation (44.4–76%) was recorded in the African Arabs who raise camels, compared to the African Arabs, who switched to the cattle farming or sedentary agriculture (1.4 to 25.5%). The results support the hypothesis formed in previous studies suggesting that the –13,915*G was at first spread in Arabia probably in connection of the camel domestication. The

estimating of the age and expansion of the $-13,915^*G$ is consistent with the migrations from Arabia to Africa which started in the 7th century AD.

Obsah

1. Úvod	4
2. Cíle práce	5
3. Materiál a metody	6
4. Výsledky a diskuze	7
5. Závěry	9
1. Introduction.....	11
2. Aims of the study.....	12
3. Material and methods	13
4. Results and discussion.....	14
5. Conclusions.....	17
Použitá literatura / References	18
Curriculum vitae.....	20
Seznam publikací / Selected publications.....	22

1. Úvod

V období kdy jsou savci kojeni, je v jejich tenkém střevě produkován enzym laktáza, který umožňuje štěpit mléčný cukr laktózu na galaktózu a glukózu. Po odstavení produkce laktázy ustává. Konzumace většího množství čerstvého mléka pak může způsobovat trávicí obtíže, neboť nestrávenou laktázu štěpí bakterie tlustého střeva a jejich metabolity mohou narušit osmotickou rovnováhu střeva. Toto se netýká konzumace mléčných výrobků, které prošly fermentací, protože laktóza je v nich tímto procesem již rozložena. U některých lidských populací se rozšířila schopnost produkovat laktázu během celého života; tzv. laktázová persistence (LP) (Swallow 2003). LP není v dnešním světě rozmístěna rovnoměrně, její nevyšší frekvence nacházíme v severní Evropě a u pasteveckých národů aridních oblastí jako je sahel nebo Arabský poloostrov. Naopak v jižní Asii či u původních obyvatel Ameriky a Austrálie se nevyskytuje vůbec (Itan et al. 2010).

Laktázová persistence je geneticky podmíněný znak. Její příčinou je jednobodová mutace ve třináctém intronu genu *MCM6* (*minichromosome maintenance gene 6*), který sousedí s genem pro laktázu. Dnes známe pět různých polymorfismů, které za LP zodpovídají a zřejmě se objevily nezávisle na sobě u různých populací. Jedná se -13910^*T (číslo označuje vzdálenost v bp *upstream* od genu pro laktázu) která je příčinou LP v Evropě (Enattah et al. 2002), -13915^*G , jejíž nejvyšší frekvence nacházíme na Arabském poloostrově (Enattah et al. 2008; Imtiaz et al. 2007) a dále pak -13907^*G , -14009^*G a -14010^*C , které se nejčastěji vyskytují ve

východní Africe (Jones et al. 2015; Ranciaro et al. 2014; Tishkoff et al. 2007).

Vznik a rozšíření LP je spjat se vznikem zemědělství před 10 000 lety na Předním východě (Bar-Yosef 1998) a s počátky mléčné produkce. LP se zřejmě rozšířila u populací, které ve své historii tradičně konzumovaly nebo stále konzumují nefermentované mléko (Holden and Mace 1997; McCracken 1971; Simoons 1970). LP je rovněž příkladem silné pozitivní selekce, která na její rozšíření v některých populacích působila (Bersaglieri et al. 2004; Enattah et al. 2008; Tishkoff et al. 2007). V severní Evropě mohl být tento selekční tlak zapříčiněn tím, že příjem většího množství mléka pomáhá suplovat absenci vápníku, která může být ve vyšších zeměpisných šířkách způsobena nedostatkem slunečního osvětlení a tím pádem nedostatkem vitamínu D získaného z prekurzorů v kůži. Vitamin D totiž hraje významnou roli v metabolismu vápníku a avitaminóza může být příčinou křivice a osteomalacie. V aridních oblastech mohlo být pak mléko výhodným zdrojem tekutin a kalorií, čímž se schopnost přijímat jeho větší množství bez obtíží mohla stát výhodnou (Flatz and Rotthauwe 1973).

2. Cíle práce

Výzkumná část této práce byla zaměřena na LP v jižní Arábii a u afrických Arabů. Na Arabském poloostrově je nejčastější příčinou LP mutace -13915*G a podle Enattah et al. 2008 a Imtiaz et al. 2007 zde mohla i vzniknout a poprvé se rozšířit, pravděpodobně v souvislosti s domestikací velblouda a jeho

využíváním pro mléčnou produkci (Enattah et al. 2008). Cíle práce jsou shrnuty do tří bodů:

1. Odhadnout alelickou frekvenci všech pro-LP mutací v Arábii a u afrických Arabů. Zjistit zda výsledky budou odpovídat dosavadní představě, že nejčastější variantou v Arábii je -13915*G, u níž se předpokládá, že se i zde poprvé objevila (Enattah et al. 2008; Imtiaz et al. 2007)
2. Porovnat subsistenci a výskyt pro-LP mutací. Ověřit předpoklad, že vznik a šíření -13915*G souvisí s domestikací velblouda a jeho šířením do Afriky kočovnými beduíny.
3. Určit dobu, kdy varianta -13915*G expandovala v Arábii a u afrických Arabů, porovnat výsledky s historickými doklady.

3. Materiál a metody

Při analýzách bylo použito celkem 620 vzorků z třinácti populačních souborů, šesti souborů z jižní Arábie – Jemenu a Ománu a sedmi souborů afrických Arabů ze Súdánu, Egypta, Čadu a Nigérie. Všechny vzorkované populace patří do semitské větve afroasijské jazykové rodiny. Zastoupeny byly jak populace usedle žijících zemědělců, tak nomádů – pastevců velbloudů a pastevců skotu.

K metodickým postupům patřila izolace DNA třemi různými metodami v závislosti na typu sběru (bukální stěry, komerční kity Oragene a Norgen). Dále PCR amplifikace a sekvenace 359 bp úseku jež obsahuje všech pět známých pro-LP variant (Coelho et al. 2009). Pro zjištění doby, kdy varianta -13915*G expandovala v Arábii a u afrických Arabů, bylo

analyzována 6 SNP v 419 kb dlouhém úseku v okolí oblasti s pro-LP mutacemi (Heyer et al. 2011), k analýzám byla použita metoda kvantitativní PCR – vysokorozlišovací analýza křivek tání. Přičemž primery byly navrženy a protokoly optimalizovány autorkou práce.

K vyhodnocení výsledků týkajících se výskytu pro-LP mutací byl použit program GeneAlex (Peakall and Smouse 2012) k určení stáří expanze mutace pak metoda spojující odhad stáří alely a rychlosti růstu (Austerlitz et al. 2003), přičemž haplotypy byly učeny programem PHASE 2.1 (Stephens and Scheet 2005; Stephens et al. 2001).

4. Výsledky a diskuze

U všech zkoumaných souborů byla identifikována alela -13915*G. Nejvyšší frekvence byla určena u Arabů Rašajda (76 %) a u Ománců ze Zufáru (65,7 %) a dále u souborů západního Jemenu (59–65,1 %). U afrických Arabů, mimo zmiňovaných Arabů Rašajda, byly nejvyšší frekvence této mutace určeny u Arabů Abbala z Čadu (44,4 %), Arabů Baggara z Čadu (19,8 %) a Arabů Baggara ze Súdánu (25,5 %). U většiny souborů z jižní Arábie jsou frekvence -13915*G vyšší než u souborů afrických Arabů. Výjimky tvoří zmiňovaní Arabové Rašajada ze Súdánu, a Jemenci z Hadramawtu (východní Jemen), u kterých byla frekvence -13915*G jen 25 %.

Arabové Rašajda přišli z Arábie do Afriky nedávno – uvádí se po roce 1860 (Young 1996) a vysoká frekvence -13915*G více odpovídá těm, co byly určeny na Arabském poloostrově i v rámci předchozích studií (Enattah et al. 2008;

Imtiaz et al. 2007) především v Saúdské Arábii. Co se týče výskytu jiných pro-LP mutací, u souborů jižní Arábie nacházíme jen malé nebo žádné zastoupení dalších pro-LP variant. Naproti tomu u souborů afrických Arabů, je zastoupeno více různých pro-LP mutací zároveň. Jedinou výjimku opět tvoří Arabové Rašajda, kde byla zaznamenána pouze -13915^*G .

Soubor z východojemenské oblasti Haramawt je naopak nižší frekvencí -13915^*G (25 %) bližší souborům afrických Arabů. Rovněž předchozí studie ukazují, že i skladbou mtDNA haploskupin se hadramawtský soubor podobá více subsaharským než okolním arabským populacím (Priehodová et al. 2017; Černý 2016). Vysvětlením může být i migrace afrických žen z tzv. hadramí diaspor, zakládaných v Africe i jinde (Manger 2010).

Fakt, že nevyšší alelické frekvence -13915^*G mezi soubory afrických Arabů nalézáme u pastevců velbloudů – Arabů Abbala z Čadu (44,4 %) a Rašajda ze Súdánu (76 %) může hovořit ve prospěch hypotézy, že vznik a šíření -13915^*G souvisí s domestikací velblouda (Enattah et al. 2008). U Arabů Baggara, kteří po své migraci do Afriky přešli k chovu hovězího dobytka, jsou frekvence -13915^*G nižší (19,8 % u Baggarů ze Súdánu a 25,5 % u Baggarů z Čadu).

Pro odhad věku a rychlosti růstu -13915^*G byly použity jednak sloučené soubory: sloučení západní Jemenci a Arabové Abbala sloučení s Araby Baggara, pro soubor Arabů Rašajda byl pak výpočet proveden samostatně, stejně tak pro soubor Ománců ze Zufáru. Nejstarší věk byl zjištěn u populací západního Jemenu 1 510 let (v intervalu 1 356–1 799 let) a u Arabů Abbala a Baggara 1 449 let (v intervalu 1 274–1 782

let). Nejnižší odhad stáří pak vyšel podle očekávání pro Araby Rašajda 250 let (v intervalu 219–312 let) a překvapivě také Ománce ze Zufáru 276 let (v intervalu 214–234 let).

Zjištěné výsledky mohou reflektovat historické události. V 6. století n. l. totiž došlo k rozpadu jihoarabských karavanních království a část beduínů, přepravujících svými velbloudy kadidlo, myrhu a další zboží z jihoarabských přístavů ke Středozeří po souši, přišla o svou původní obživu. Produkce a konzumace velbloudího mléka se tak pro ně stala důležitější než dříve, což se mohlo projevit i expanzí –13915*G. Od 7. století n. l. rovněž začínají výraznější migrace z Arábie do Afriky, s nimiž sem pravděpodobně, podle historických pramenů, přišli také předkové dnešních Arabů Abbala a Baggara.

Nízký věk expanze mutace –13915*G jež byl spočten u súdánských Arabů Rašajda na 250 let, se rovněž nerozchází s historickými fakty, neboť jejich předkové přišli z Arábie do Afriky po roce 1860 n. l. (Young 1996).

Zjištěné výsledky tak ukazují, jak lze díky genetické podstatě LP sledovat kromě subsistence i proběhlé migrační trasy a genový tok.

5. Závěry

1. U všech souborů jižní Arábie a afrických Arabů byla nalezena pro-LP varianta –13915*G, přičemž v Arábii byly její frekvence vyšší (56,9–65,7 %) než u afrických Arabů (1,4–44,4 %) s výjimkou souboru z jemenského Hadramawtu (25 %) a súdánských Arabů Rašajda (76 %). U afrických Arabů byl pak podíl ostatních pro-LP variant

větší než u souborů z jižní Arábie. Výsledky jsou v souladu s předchozími studii a podporují arabský původ –13915*G.

2. U afrických Arabů chovajících velbloudy nacházíme větší podíl varianty –13915*G než u afrických Arabů, kteří přešli k chovu hovězího dobytka či usedlému zemědělství. Výsledky jsou v souladu s hypotézou o vzniku a rozšíření –13915*G v souvislosti s domestikací a chovem velblouda v Arábii a následně v Africe.
3. Odhady věku expanze ukazují, že –13915*G se v západním Jemenu rozšířila před ~1 500 lety, což souvisí s rozpadem karavanních království v jižní Arábii v 6. století n. l. Díky oslabení obchodu, k němuž byli velbloudi využíváni, se jejich chov přeorientoval spíše k mléčné produkci. Vypočtená doba expanze –13915*G u Arabů Abbala a Baggara – před ~1 450 lety zřejmě poukazuje na migrace z Arábie probíhající od 7. století n. l., v rámci kterých předkové dnešních Arabů Abbala a Baggara přišli do Afriky.

1. Introduction

During the infancy, when new-borns depend on breast-feeding, in an enzyme called lactase there is produced their small intestine. Lactase cleaves the milk sugar called lactose to galactose and glucose. After the weaning, the lactase production declines. Consequently, consuming of larger amounts of fresh milk may cause digestive problems as the undigested lactase is cleaved by colonic bacteria and metabolites may cause the osmotic imbalance in the colon. Dairy products, which passed fermentation, do not provoke these problems, because the lactose there is already cleaved. In some human populations, it the ability to produce the lactase throughout life is widespread. This phenomenon is called lactase persistence (LP) (Swallow 2003). Worldwide distribution of this trait is not uniform, the highest frequencies are found in northern Europe and among pastoral tribes in arid regions in Africa and the Arabian Peninsula. By contrast, in southern Asia and among the Native Americans and the Australians LP does not occur at all (Itan et al., 2010).

Lactase persistence is a genetically determined trait, caused by a single nucleotide mutation (SNP) in the flanking region of the lactase gene – in the intron 13 of *MCM6* gene (minichromosome maintenance gene 6). Interestingly, five different SNPs are known to be responsible for LP and they apparently emerged independently in different populations. The $-13,910^*T$ mutation (the number means the distance in bp upstream of the lactase gene) causes the LP in Europe (Enattah et al. 2002), the $-13,915^*G$ is found in the highest frequency in

the Arabian Peninsula (Enattah et al. 2008; Imtiaz et al. 2007), The $-13,907^*G$, $-14,009^*G$ and $-14,010^*C$ are most common in Eastern Africa (Jones et al. 2015; Ranciaro et al. 2014; Tishkoff et al. 2007).

The expansion of LP is linked with the origins of agriculture 10,000 years ago in the Near East (Bar-Yosef 1998) and the beginnings of milk production. LP apparently spread among populations with a custom of milk production and those who have traditionally consumed and sometimes still consume unfermented milk (Holden and Mace 1997; McCracken, 1971; Simoons 1970). LP is also an example of strong positive selection, which occurred in some populations (Bersaglieri et al. 2004; Enattah et al. 2008; Tishkoff et al. 2007). In northern Europe, the selection could be caused by the advantage of improved calcium absorption. The intake of large quantities of milk helps to substitute the lack of calcium, which is in the caused due to the lack of sunshine higher latitudes. Thus the deficiency of vitamin D obtained from skin precursors may develop and because vitamin D plays an important role in the calcium metabolism, its deficiency may lead to rickets and osteomalacia. In arid areas, milk could then be the good source of fluids and calories, therefore the ability to accept its larger quantities without difficulty could be advantageous (and Flatz Rotthauwe 1973).

2. Aims of the study

The research part of this work is focused on the LP in southern Arabia and in the African Arabs. The $-13,915^*G$ is the most common cause of the LP in the Arabian Peninsula which is the

place of its origin as well (Enattah et al. 2008, Imtiaz et. al 2007). It is possible that its origin is related to the domestication of camels for milk production.

The goals of the study are summarized in three points:

1. Estimating the allelic frequencies of the LP mutations in Arabia and in the African Arabs. Determining if results correspond to the existing finding that the –13,915*G is the most common variant in Arabia, where it probably first appeared (Enattah et al. 2008; Imtiaz et al. 2007)
2. Comparing the subsistence and the incidence of the LP mutations. Verifying the assumption that the emergence and the spread of the –13,915*G is related to the domestication of camels and its spread to Africa led by the nomadic Bedouins.
3. Determining when the –13,915*G expanded in Arabia and in the African Arabs, comparing results with historical facts.

3. Material and methods

A total of 620 samples from thirteen population samples, six population samples from southern Arabia – Yemen and Oman and seven population samples of the African Arabs from Sudan, Egypt, Chad and Nigeria were used for the analyses. All sampled populations belong to the Semitic branch of the Afro-Asiatic language family. The samples consist of the populations of sedentary farmers and nomads – both camel and cattle herders.

Used methods contained the isolation of DNA using three different methods depending on the type of the collection (buccal swabs, commercial kits – Oragene, Norgen). Further, a PCR amplification and sequencing of 359 bp fragment which contains all five known LP-variants was performed (Coelho et al. 2009). To determine the time of expansion of the –13,915*G in Arabia and in the African Arabs, the 6 SNPs in 419 kb long stretch around the LP mutations were analysed (Heyer et al. 2011). A method of quantitative PCR called “High resolution melting” was used for this analysis. Primers were designed and protocols were optimised by the author.

Results concerning the incidence of the LP mutations were evaluated using the GeneAlex software (Peakall and Smouse 2012). The age of expansion of the mutation was determined by the method joining estimation of the allele age and its growth rate (Austerlitz et al., 2003). Haplotypes were estimated by the program PHASE 2.1 (Stephens and Scheet 2005; Stephens et al. 2001).

4. Results and discussion

The –13,915*G was identified in all the population samples. The highest frequency was found in the Rahaayda Arabs (76%) and the Omanis from Dhofar region (65.7%), followed by the western Yemenis (59–65.1%). The highest frequency of this mutation in the African Arabs (except of the mentioned Rahaayda) were found in the Abbala from Chad (44.4%), the Baggara from Chad (19.8%) and the Baggara from Sudan (25.5%). The frequencies of the –13,915*G were higher in most samples from southern Arabia than in the samples of

African Arabs. The exceptions are the mentioned Rashaayda Arabs from Sudan, and the Yemenis from Hadramawt (eastern Yemen), where the frequency of $-13,915^*G$ is only 25%.

The Rashaayda Arabs came from Arabia to Africa recently, only after 1860 (Young, 1996). High frequency of the $-13,915^*G$ found here corresponds more to frequencies in the Arabian Peninsula, also reported by previous studies (Enattah et al. 2008; Imtiaz et al. 2007). Regarding the incidence of the other LP mutations in South Arabia, it was found little or absence of the other LP variants. In contrast, several LP mutations were represented simultaneously in the population of African Arabs. The only exception is again the Rashaayda Arabs with only the $-13,915^*G$. The eastern Yemeni Haramawt are closer to the African Arabs due to lower frequency of $-13,915^*G$ (25%). Also, previous studies showed that even a composition of Haramawt's mtDNA haplogroups resembles rather to sub-Saharan populations than surrounding Arabic populations (Priehodová et al. 2017; Černý 2016). This can be explained by immigration of African women from so-called Hadrami diasporas which were set up in Africa and elsewhere (Manger 2010).

The highest allelic frequency of $-13,915^*G$ of the African Arabs was identified in the Abbala from Chad (44.4%) and the Rashaayda from Sudan (76%). These two populations are camel herders. Such results correspond to the hypothesis that the emergence and spread of the $-13,915^*G$ is related to the domestication of camels (Enattah et al. 2008). The frequencies of the $-13,915^*G$ in Baggara Arabs are lower (19.8% in the Sudanese Baggara and 25.5% in the Chadic Baggara) and are

explained via switching to the cattle herding after their migration to Africa.

In order to estimate the age and the growth rate of the $-13,915^*G$, merged population samples were used: the merged western Yemenis and the Abbala Arabs merged with the Baggara Arabs. For Rashayda Arabs the calculation was performed separately, as well as for the Omanis from Dhofar. The oldest age was determined in western Yemen 1,510 YBP (range 1,356–1,799 YBP) and in the Abbala and Baggara Arabs 1,449 YBP (range 1,274–1,782 YBP). The lowest age was obtained for the Arabs Rashaayda 250 YBP (range 219–312 YBP) and the Omanis from Dhofar 276 YBP (range 214–234 YBP). These results reflect the historical events. In the 6th century the Southern Arabian Caravan Kingdoms collapsed and part of the Bedouins that used their camels for transport of frankincense and myrrh from southern Arabia, have lost their original livelihood. Therefore, the production and the consumption of camel milk may have started to be more important than before and is reflect in the expansion of the $-13,915^*G$. The migrations from Arabia to Africa have begun from the 7th century AD, within which the ancestors of today's Abbala and Baggara Arabs probably came. Early age of the expansion of the mutation $-13,915^*G$ obtained in the Sudanese Rashaayda Arabs is only 250 YBP, which corresponds to historical facts, since their ancestors came from Arabia to Africa after 1860 AD (Young 1996).

Results show that the research of the LP provide not only information about the population subsistence, but also about the past migration routes and gene flow.

5. Conclusions

1. The LP mutation $-13,915^*G$ was found in all the population samples from southern Arabia and the African Arabs. The samples from Arabia have higher frequency (56.9 to 65.7%) than African Arabs samples (1.4 to 44.4%), with the exception of the Yemenis Hadramawt sample (25%) and the Sudanese Rashaayda Arabs (76%). For the African Arabs was the proportion of other LP variants larger than for the population samples from southern Arabia. Results are consistent with previous studies and support Arab origin of the mutation $-13,915^*G$.
2. Higher frequency of the $-13,915^*G$ was found in the African Arabs, who are camel herders, rather than in the African Arabs, who switched to cattle herding or sedentary agriculture. Results are in agreement with the hypothesis of the origin and expansion of the $-13,915^*G$ is connected with the camel domestication in Arabia.
3. The age estimation shows that the $-13,915^*G$ in western Yemen spread $\sim 1,500$ years ago, which may be associated with the collapse of the Southern Arabian Caravan Kingdoms in the 6th century AD. The caravan trade based on camel workforce weakened and the camel herders switched more to dairy production. The obtained age of the expansion of the $-13,915^*G$ in the Abbala and the Baggara Arabs is $\sim 1,450$ years ago, this date is consistent with migrations from Arabia since the 7th century A.D. when ancestors of today's Baggara and Abbala Arabs arrived in Africa.

Použitá literatura / References

- Austerlitz F, Kalaydjieva L, and Heyer E. 2003. Detecting population growth, selection and inherited fertility from haplotypic data in humans. *Genetics* 165(3):1579-1586.
- Bar-Yosef O. 1998. The Natufian culture in the Levant, threshold to the origins of agriculture. *Evolutionary Anthropology: Issues, News, and Reviews* 6(5):159-177.
- Bersaglieri T, Sabeti PC, Patterson N, Vanderploeg T, Schaffner SF, Drake JA, Rhodes M, Reich DE, and Hirschhorn JN. 2004. Genetic signatures of strong recent positive selection at the lactase gene. *Am J Hum Genet* 74(6):1111-1120.
- Coelho M, Sequeira F, Luiselli D, Belezza S, and Rocha J. 2009. On the edge of Bantu expansions: mtDNA, Y chromosome and lactase persistence genetic variation in southwestern Angola. *BMC Evol Biol* 9:80.
- Černý V. 2016. Po stopách 'Ádů. Praha: Academia.
- Enattah N, Jensen T, Nielsen M, Lewinski R, Kuokkanen M, Rasinpera H, El-Shanti H, Seo J, Alifrangis M, Khalil I et al. . 2008. Independent introduction of two lactase-persistence alleles into human populations reflects different history of adaptation to milk culture. *Am J Hum Genet* 82(1):57-72.
- Enattah NS, Sahi T, Savilahti E, Terwilliger JD, Peltonen L, and Järvelä I. 2002. Identification of a variant associated with adult-type hypolactasia. *Nat Genet* 30(2):233-237.
- Flatz G, and Rotthauwe HW. 1973. Lactose nutrition and natural selection. *Lancet* 2(7820):76-77.
- Heyer E, Brazier L, Ségurel L, Hegay T, Austerlitz F, Quintana-Murci L, Georges M, Pasquet P, and Veuille M. 2011. Lactase persistence in central Asia: phenotype, genotype, and evolution. *Hum Biol* 83(3):379-392.
- Holden C, and Mace R. 1997. Phylogenetic analysis of the evolution of lactose digestion in adults. *Hum Biol* 69(5):605-628.
- Imtiaz F, Savilahti E, Sarnesto A, Trabzuni D, Al-Kahtani K, Kagevi I, Rashed MS, Meyer BF, and Järvelä I. 2007. The T/G 13915 variant upstream of the lactase gene (LCT) is the founder allele of lactase persistence in an urban Saudi population. *J Med Genet* 44(10):e89.
- Itan Y, Jones BL, Ingram CJ, Swallow DM, and Thomas MG. 2010. A worldwide correlation of lactase persistence phenotype and genotypes. *BMC Evol Biol* 10:36.

- Jones BL, Oljira T, Liebert A, Zmarz P, Montalva N, Tarekeyn A, Ekong R, Thomas MG, Bekele E, Bradman N et al. . 2015. Diversity of lactase persistence in African milk drinkers. *Hum Genet* 134(8):917-925.
- Manger LO. 2010. *The Hadrami diaspora: community-building on the Indian Ocean rim*. New York: Berghahn Books.
- McCracken RD. 1971. Lactase deficiency - example of dietary evolution. *Current Anthropology* 12(45):479-497.
- Peakall R, and Smouse PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics* 28(19):2537-2539.
- Priehodová E, Austerlitz F, Čížková M, Mokhtar MG, Poloni ES, and Černý V. 2017. The historical spread of Arabian Pastoralists to the eastern African Sahel evidenced by the lactase persistence -13,915*G allele and mitochondrial DNA. *American Journal of Human Biology*:e22950-n/a.
- Ranciaro A, Campbell MC, Hirbo JB, Ko WY, Froment A, Anagnostou P, Kotze MJ, Ibrahim M, Nyambo T, Omar SA et al. . 2014. Genetic origins of lactase persistence and the spread of pastoralism in Africa. *Am J Hum Genet* 94(4):496-510.
- Simoons FJ. 1970. Primary adult lactose intolerance and the milking habit: a problem in biologic and cultural interrelations. II. A culture historical hypothesis. *Am J Dig Dis* 15(8):695-710.
- Stephens M, and Scheet P. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am J Hum Genet* 76(3):449-462.
- Stephens M, Smith NJ, and Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68(4):978-989.
- Swallow DM. 2003. Genetics of lactase persistence and lactose intolerance. *Annu Rev Genet* 37:197-219.
- Tishkoff SA, Reed FA, Ranciaro A, Voight BF, Babbitt CC, Silverman JS, Powell K, Mortensen HM, Hirbo JB, Osman M et al. . 2007. Convergent adaptation of human lactase persistence in Africa and Europe. *Nat Genet* 39(1):31-40.
- Young WC. 1996. *The Rashaayda Bedouin: Arab Pastoralists of Eastern Sudan*: Harcourt Brace College Publishers.

Curriculum vitae

Adresa / Address: Vašírov 1, 270 61 Lány Vašírov
Datum narození / Born: 18. 7. 1985
Kontakt / Contact: priehodovaedita@seznam.cz;
+420 608 669 911

Vzdělání / Education:

2009: Bc. Biologie PřF UK /

Bachelor degree Biology, Charles University

2011: Mgr., Antropologie a genetika člověka, Katedra antropologie a genetiky člověka, PřF UK /

Master degree, Anthropology and human genetics, Charles University

2011–současnost / –today:

Ph. D. program Antropologie a genetika člověka, Katedra antropologie a genetiky člověka, PřF UK /
Ph. D. program Anthropology and human genetics Charles University

7. 4. 2015:

SZZ Antropologie a genetika člověka /
Final doctoral exam Anthropology and human genetics

Stáže / Scholarships:

3. 1. 2011–17. 6. 2011:

Erasmus, Université Bordeaux 1

13. 9. 2013–2. 12. 2013:

French government's scholarship program, Université Paris Diderot / Musée de l'Homme, Paris

Odborné zkušenosti / Working experience

2010–současnost / –today:

Částečný pracovní úvazek v Laboratoři archeogenetiky,
Archeologický ústav AV ČR, Praha, v. v. i. /
Institute of Archaeology of the CAS, Prague, V. V. I.

Granty / Grants

2012–2014:

Laktázová persistence a její asociace k subsistenční strategii: příspěvek k populační historii afrického sahelu (GAUK 651112), hlavní řešitelka grantu /
Lactase persistence and association with life style: contribution to the population history of the African Sahel (GAUK 651112); principal investigator

2011–2013:

Chronologické a geografické perspektivy v genetické diverzitě nerekombinantních lokusů z Česka a Slovenska, spoluřešitelka /
Chronological and geographical perspectives in genetic diversity of non-recombinant loci in Czechia and Slovakia (GAUK 270911), co-investigator

Jazyky / Languages

Anglický jazyk – středně pokročilý (FCE certifikát) /
English: upper-intermediate (FCE certificate)

Francouzský jazyk – středně pokročilý /
French – upper-intermediate

Německý jazyk – začátečník /
German – beginner

Informatika / Informatics skills

UNIX, R

Seznam publikací / Selected publications

Publikace, na nichž je disertační práce založena / Publications on which dissertation is based

Priehodová E, Abdelsawy A, Heyer E, and Cerný V. 2014. Lactase persistence variants in arabia and in the african arabs. Hum Biol 86(1):7-18.

Priehodová E, Austerlitz F, Čížková M, Mokhtar MG, Poloni ES, and Černý V. 2017. The historical spread of Arabian Pastoralists to the eastern African Sahel evidenced by the lactase persistence -13,915*G allele and mitochondrial DNA. American Journal of Human Biology:e22950-n/a..

Populárně naučné / popularisation of science

Priehodová E. 2013. Pravěk ve střední Evropě – svědectví mitochondriální DNA. Živa 1:4-6.

Priehodová E. 2016. Laktázová persistence a pití mléka. Živa 5:238-240.