**Title of the Satellite Meeting:** Local History and Genealogy Satellite Meeting 2019  
**Date:** 21-22 August 2019  
**Location:** University library “Svetozar Marković”, Belgrade, Serbia

Digital Library in The Service of Genetic Genealogy

**Velibor Mihich**  
Committee for Digitization, Society of Serbian Genealogists “Poreklo”, Belgrade, Serbia.  
E-mail address: velibor.mihic@poreklo.rs

**Milan Rajevac**  
Editorial Board of the Serbian DNA project, Society of Serbian Genealogists “Poreklo”, Belgrade, Serbia.  
E-mail address: milan.rajevac@poreklo.rs

**Branko Todorovich**  
Committee for Digitization, Society of Serbian Genealogists “Poreklo”, Belgrade, Serbia.  
E-mail address: branko.todorovic@poreklo.rs

---

Copyright © 2019 by Velibor Mihich, Milan Rajevac, Branko Todorovich. This work is made available under the terms of the Creative Commons Attribution 4.0 International License:  
http://creativecommons.org/licenses/by/4.0

---

**Abstract:**

In the research of origin, genetic genealogy allows us to go to the earliest of our ancestors. However, the results of DNA analysis alone are not enough to establish and explain the relationships among the individuals. For more precise conclusions, it is necessary to be examined what has been recorded in the historical sources and literature. The Society of Serbian Genealogists “Poreklo” (“Origin”), which started Serbian DNA project in 2012, several years later also started work on forming special Digital Library for genealogical research. From that time, project grew covering many geographical regions of Western Balkans and ever-increasing Y-DNA database. Digital library also grew and now contains more than 3,800 titles of genealogical character, covering many regions of Western Balkan and even territories tied to historical migration routes out of this region. The special quality of this library is a convenient word or phrase search of all digitized books. Consequently, every new genetic result which contains the main information about a tested person can be
checked in one simple and fast way. In many cases, this helps to figure out the secrets of genetic origins.

Keywords: Digital Library, Serbian DNA project, Genetic genealogy, Western Balkan.

---

**Introduction**

The beginning of genetic genealogy date back to the end of the 19th century, but it did not become popular until the beginning of the 21st century, when Bryan Sykes, a molecular biologist at Oxford University, published his study “Surnames and the Y-Chromosome” (2000). He tested the new methodology in general surname research, based on the tracking of four STR markers on the Y-chromosome. The Y-DNA is transmitted exclusively from father to son, pointing to a significant potential that genetics could have in genealogy and history.

Genetic genealogy is one of the most important auxiliary historical sciences today, experiencing its greatest use in the domain of family origin research. It has become an indispensable part of any serious genealogy research. With an increase in Y-DNA testing, the number of STR markers is now to hundreds.

The results obtained by genetic testing alone are not enough to explain somebody’s origin. It is of great importance to not only investigate family tradition, particular sources, and archival sources but to also explore relevant genealogical literature. As we are talking of numerous, various and often unsystematic written literature, the best model for fast large-scale research showed up to be forming of Digital Library, which in our case will be based on Rescarta’s software platform.

---

1 It is a software package for management and distribution of digital content that provides simple and unified way of documents reading, based on standard WEB browsers. ResCarta can be also used for storing of digital audio recordings as well as images, but for the purpose of Poreklo’s digital library (http://biblioteka.poreklo.rs) it is utilized only for saving digital books, journals and papers.

Using ResCarta it is possible to organize and classify units from digital library collection in different thematic groups, giving different access rights to users, depending of their privileges. ResCarta server is implemented on Apache HTTP server platform. All tools from ResCarta software package can be used free of charge, as open source under Apache 2.0 license.

Before uploading on ResCarta server, digitized material goes through a process of digital processing and graphical data format adjustment. All library items are uploaded on ResCarta server as collections of digital images in one of standard raster (bitmap) graphic formats, like JPEG, PNG or TIF. As its consequence, all documents originally recorded in object-stored formats, like digitally created PDFs or DOCs, should be converted to raster images.

Very important step in digital pre-processing of scanned material is optimization of graphical image format. Image scanners and digital cameras used for digitalization of documents usually store images in full 24-bit color format. Digitally stored scanned copy of an average book in the full color format often occupies few hundred megabytes. For a large book collection of several thousand units it requires significant computer hard disk capacities and slows down data processing during regular digital library usage. However, in most cases,
Serbian DNA project & Digital Library “Poreklo”

This is exactly what the Society of Serbian Genealogists “Poreklo” did. The Serbian DNA project was started in 2012, and day by day its database grew. It now contains about 5000 Y-DNA results. Already then, it was clear how important it is to have the most important genealogical literature in digital and researchable format. The process of digitization has since begun and the researchable Digital Library “Poreklo was launched at the end of 2016.

This was a big step forward, especially in terms of fast research capabilities. This challenge can be seen as classical problems of research. Best case scenario, there is an index at the end of the book. In most cases there isn’t more than the content of the book, in which case you must read page by page, searching for the term or part that you’re interested in. For someone’s individual research, it may not be that such an insurmountable problem, but when you have to research different surnames every day and find something within a short period of time, the researchable digital library successfully prevails that problem.

In order to best utilize the results determined by the laboratory, for the further course of research, candidates are asked to fill out a questionnaire before Y-DNA testing. In the questionnaire, they provide basic information about their origin and three significant data there is no need for storing digitized material in full color format, as most of it is consisted of pure textual black and white pages. Conversion of such pages to monochromatic 1-bit black and white graphical format reduce size of digital footprint of scanned items for ten times or even more, without any significant loss on the image quality.

Conversion of digitalized items to the monochromatic graphical format may have other benefits. ResCarta is using own built-in algorithm for optical character recognition (OCR) which is used for indexing and achieving searchability of the digital library content. The OCR algorithm shows much better results for a text written with clear fonts, and that effect can appear with conversion to monochromatic color format.

For pre-processing of digitized documents a set of applications is used – Adobe Acrobat Pro, ScanTailor, GIMP 2.0 and IrfanView.

Let’s also mention that the digital library catalog of publications (https://katalog.poreklo.rs), which was developed as a separate VEB platform, was built on the Xataface software package (http://xataface.com) as an auxiliary tool for more comfortable use of the Digital Library “Poreklo”. The digital catalog contains a list of all the library units, as well as individual articles in it, and enables a very simple and efficient search of the library by authors, titles, publishers, publishing years, and the like. Each catalog unit is hyperlinked with the corresponding library unit, allowing direct access to the requested material.

Serbian DNA project (https://dnk.poreklo.rs) was established in September 2012 and represents the pioneer venture of enthusiasts gathered around Society of Serbian Genealogists “Poreklo” to gather information about all tested people from our region who have done Y-DNA test. The database of Serbian DNA project contains results of tested Serbs and other genetically close individuals from the territory of Serbia, Former Yugoslavia and other neighbor countries. Those are the results of individuals tested in laboratories with which they cooperate, as well as those that are taken from the public databases of international DNA projects. The aim of Serbian DNA project is to discover the secrets of origin through the genetic genealogy and to try to make new insights into our history.
points: *Surname, Krsna slava* (Family Saint-Patron) and the place where their ancestor lived in 1900. Here, the Krsna slava especially stands out, because it is the characteristic of Orthodox Serbs, who make up the largest number of Western Balkans tested individuals in Serbian DNA project. It is custom to celebrate a saint or family patron, which is usually transferred from father to son, as well as the surname and the Y-chromosome markers. Bearing in mind that the Serbs celebrate more than 70 different saints during the year, it comes out to be another parameter in addition to the surname and place of origin. Therefore, the combination of these data points helps us determine the origin of the tested individual, especially when we are talking about those frequent surnames or places which contain a large number of families.

The results of the tested Y-DNA are presented by a series of numbers - markers (12, 17, 23, 37, 64, 111, 500, 700), that help us determine the haplogroup and its subclade. For this purpose, we use special software called Y-DNA Predictor³.

That is the moment where we activate the searchable digital library. Through this, and by combining the three previously mentioned pieces of information, we narrow in on the relevant literature and important parts, which tell us about the subject’s family origin. On this basis, his results will be placed in a certain genus in the database of the Serbian DNA project.

However, the genus is not always historically and anthropo-geographically clearly defined, i.e. apart from the fact that a certain group of families is genetically related, often, there is no other common denominator, through which we could place that genetic genus in some clearer historical context. In the beginning, this was the case for most of the tested. Clans primarily arose from the territories of Montenegro and Old Herzegovina, which in this form survived until the time of early 1900s. They were at large described in ethnological and anthropo-geographical literature, so by testing different families and brotherhoods of the clan, very quickly it became clear which clan (which origin, both traditionally and in literature, was based on the common male ancestor) belongs to which haplogroup. In some cases, those who were tested were just expecting to confirm their family origin. Sometimes the tested one didn’t have any family tradition, but after testing he became part of some genetic group. Upon comparison with others in that group, sometimes it became also easier to find the right family in the literature. It is not rare to search for a family with common surname, bumping you in many different directions. What is important here is the fact that sometimes you find the same surname in the place or area of origin of tested one, even when the tested one didn’t know much or didn’t know anything about the place or area of his origin. There is one more interesting thing here, that is the family’s saint-patron, which in some cases shows to be the same among the members of one genetic group of families, who don’t have traditions about their common origin. This was the case of a group of families that celebrate St. Emperor

---

³ Y-DNA predictor is a special program that determines haplogroup of the tested one, using delivered markers. Currently the best Y-DNA predictor in World is the NevGen program ([http://nevgen.org](http://nevgen.org)), devised and developed by Aleksandar Nevsky, and adjusted for internet by Milos Cetkovic Gentula, both the members of “Poreklo”.
Constantine and Empress Helena. They were tested separately and didn’t know that they are related to other families that celebrate the same family saint-patron.

The greatest surprise in our several years of research was the discovery of a clan that was believed to be extinct. This finding was the combination of genetic results and a significant number of ethnological and anthropo-geographical literature that was already available in the Digital Library. It is mentioned in sources that date back to as early as the 13th century. It is about Krici, a pre-Slavic population organized in a clan, which according to tradition lived in the area of Nevesinje (Bosnia and Herzegovina) in the time of Slavice movement to the Balkans. They entered into conflict with new populations and soon later were pushed to north-east in the area of Potarje (Montenegro)

The first tested one, who was determined to belong to the haplogroup J2b1-M205, which will later be known as a haplogroup of the lost clan Krici, was a member of family Mijovic (сур. Мијовић) from Годијелji (сур. Годијељи). This village belonged to the Municipality of Savnik (Montenegro), and the family saint-patron was St. Thomas (сур. Томиндан). After we entered these terms in researchable Digital Library, the software found the book Drobnjak : families in Drobnjak and their origin4 (eng. translation), where on one page we had a full match, while reducing the number of researching terms for one (we have exclude the family saint-patron), the number of results raised up to six library units, and where all of the results were referring to the family Mijovic from Godijelji. From this we could see that by reducing the number of research terms, number of results grow, in this case, bringing us the whole series of other relevant literature.

Here we found out that old surname of Mijovic was Godijelj, just like the place where they originated from. On this basis, we can conclude that the place was named after the family.

Not long after Mijovic, Pejanovc was also tested. Originally from the same place, and with the same family saint-patron, the same genetic results were obtained. Through the same research method, we have found out that the Pejanovic family is a branch of wider genus of Godijelji5. Therefore, this genetically related group of families was at first named Godijelji. However, we still didn’t know much about the earlier origin of these two families, especially because of their different family traditions. One told of the origin being Katunska Nahia (Montenegro), and other suggested Crmnica (Montenegro), which were in contradiction. This was the result of an increasing number of tested members of this haplogroup. Neither group was favored, and in those two regions there were no results like theirs.

5 Same.
Then came the result of Klickovic (сиг. Кличковић), originally from Lacarak, Municipality of Sremska Mitrovica, with St. George (сиг. Ђурђевдан) as family saint-patron. Colonized in Srem, they had family tradition originating in Dugo Selo (сиг. Дуго Село) near Vrginmost (Croatia). By searching terms “Kličković” and “Dugo Selo” (this time in Latin, since the Cyrillic search didn’t make any results), several results showed up in Digital Library. From this the most important information came from the List of victims of war 1941-1945 for Croatia⁶ (eng. translation), from which we could confirm that surname Klickovic was present in Dugo Selo at that time (thirteen of them were recorded as victims of war). However, there were no earlier mentions of them in our results, until we began to search the history of Dugo Selo. We found the book Memorial to Holy Serbian Warriors 1912-1918: Overseas Lands⁷, where in the long list of war veterans born in Dugo Selo there was one person with surname Krickovic (сиг. Кричковић). That led us to doubt the possibility that Krickovic was an older version of surname Klickovic. Also this was the first trace that pointed to the possible relationship between genus Godijelji and old clan of Krici, which lived on the same area mainland of Godijelji.

Soon after that, we tested the candidate who had root Kric- in his surname, and that was Kricka (сиг. Кричка) from the village Kricke (сиг. Кричке), Municipality of Drnis (Croatia), with St. Nicholas (сиг. Никољдан) as his family saint-patron. The early mention of this family goes back to the mid-16th century when they were most-likely also settled there. The name of village first was mentioned in 1596. Nevertheless, what is most important here, it is the result that we came through with testing, which coincided with the previously tested Godijelji. By additional checks, i.e. by testing several other families who had in their surname root Kric- (e.g. Kricak, Krickovic), and who didn’t have family traditions of their common origin, we have confirmed that all of them belong to the same haplogroup J2b1-M205. From this, we came to significant discoveries in the genetics of an old clan, previously believed to be extinct, as well as to the confirmation of historical sources that are somewhat older Balkan population than Slavs. This was separately concluded by comparing the results of deep testings completed by several members of this genetic group.

By this time, more than 200 candidates belonging to the haplogroup of Krici were tested in the Serbian DNA project. The literature in the Digital Library “Poreklo” counts 199 Cyrilic and 36 Latin-based highly significant library units, which could be used in further research regarding Krici origin, as well as for the origin of individual families that draw from this old Balkan clan.

---


⁷ Петровић, И. (2005). Поменик светих српских ратника (1912-1918) : Прекоморске земље. Нови Сад : Ауторово издање
Conclusion

In this example, we could see how one thematic digital library, like Digital Library “Poreklo”, can be of great use for interpretation of Y-DNA results. This is especially true in terms of regions that are well covered with anthropo-geographical literature. And of course, this is in addition to the representative sample of tested individuals, who in addition on their results, bring us the basic information on their nearer origin, creating sufficient basis for the deeper classic genealogical researches.

Note

The Digital Library “Poreklo”, as well as to the genetical database of the Serbian DNA project, is service with limited access. Full access is reserved only for the editorial board of the Serbian DNA project, and for the researchers in the Society of Serbian Genealogists “Poreklo”, who engage in research work on the topics of classic and genetic genealogy.

Images

![Image 1. The web-site of Serbian DNA project](image-url)
Image 2. The web-site of Digital Library “Poreklo”

Image 3. Search board in Serbian DNA project

Image 4. The results of tested individuals founded by searching terms “Мијовић” (Cyr. Мијовић), “Ст. Томан” (Cyr. Томиндан) and “Годијељи” (Cyr. Годијељи) in Cyrillic
Image 5. The results shown on the Google map

Image 6. The searching board in Digital Library “Poreklo”

Image 7. The results of literature founded by searching the terms “Mijovic” (Сур. Мијовић), “St. Thomas” (Сур. Томиндан) and “Godijelji” (Сур. Годијељи) in Cyrillic
Image 8. The content from the book founded by searching the terms “Mijovic” (Cyr. Мијовић), “St. Thomas” (Cyr. Томиндан) and “Godijelji” (Cyr. Годијељи) in Cyrillic.

Image 9. The content from the book founded by searching the terms “Kličković” and “Dugo Selo” in Latin.

Image 10. The content from the book founded by searching the terms “Kričković” (Cyr. Кричковић) and “Dugo Selo” (Cyr. Дуго Село) in Cyrillic.
Image 11. The number of tested individuals in Serbian DNA project belonging to J2-M205 haplogroup