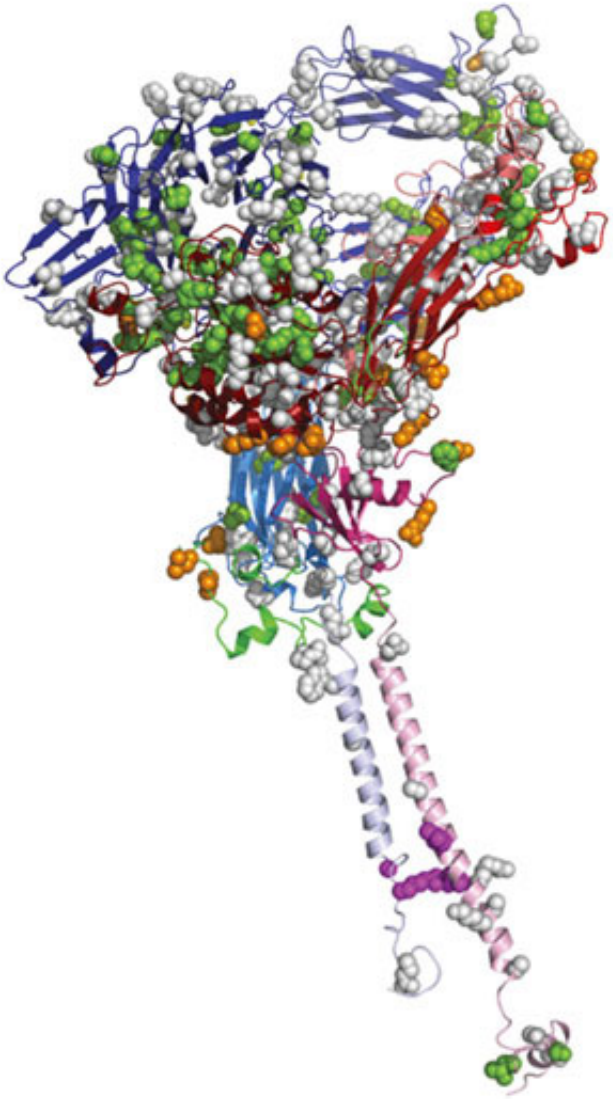
Blood disorder study illustrates the challenges to parsing genetic data

April 14 2015



Researchers pinpointed the locations on the blood platelet receptor proteins

affected by DNA variations that alter the receptor's amino acid building blocks, including those associated with the bleeding disorder Glanzmann thrombasthenia (green circles), those newly identified in the study (grey), and with other conditions (orange and magenta). Credit: Laboratory of Blood and Vascular Biology at The Rockefeller University/*Proceedings of the National Academy of Sciences*

Accumulating data, even genetic data, is easy. Understanding the meaning of those data can be more of a challenge. As genetic testing becomes increasingly popular, more and more patients and physicians are faced with tough questions: Does a particular genetic variation translate into a predisposition to an illness, or is it simply a benign rearrangement of letters with no immediate health impacts?

To try to make sense of all the genomic data being generated, researchers at Rockefeller University studied the rate and impact of mutations in two genes associated with a blood disorder. The research was published March 31 in the *Proceedings of the National Academy of Sciences*, and reflects a collaboration with scientists at the Icahn School of Medicine at Mount Sinai and the ThromboGenomics consortium.

Scanning genomic data from thousands of people, the scientists—led by postdoctoral fellow Lorena Buitrago, director of research bioinformatics Yupu Liang, and Barry S. Coller, physician in chief and vice president for medical affairs at Rockefeller—found that about 1.3 percent of the approximately 16,000 individuals analyzed carry a mutation in at least one of the two genes they studied, which encode the proteins that make up a key receptor in blood clotting. Individuals with deleterious mutations in both copies of either gene have a rare disorder of their blood platelets, the cell fragments that prevent bleeding, termed Glanzmann thrombasthenia. Roughly 10 percent of the amino acid

building blocks in these proteins were affected by the discovered mutations.

The authors then used three different commonly used algorithms to try to predict the proportion of these DNA variants that are likely to have a negative impact on health by causing excess bleeding. They got a wide range of results: Depending on the algorithm and other variables, between 27 and 71 percent of mutations were predicted to be harmful, the authors report.

To test the validity of the predictions, Buitrago made three of the variants in a cell line and examined whether they affected the production or function of the receptor. Two variants that the algorithms predicted would be deleterious did, indeed, severely affect production of the receptor. For the third variant, the algorithms were split on whether it was deleterious—in Buitrago's experiment, this variant caused a partial decrease in the production of the receptor, but did not harm its function. "In essence, the algorithms collectively got the right answer," says Buitrago.

"It was also really striking that, in such a large population, we didn't find any of the more than 100 previously reported disease-causing mutations in these two genes," says Coller. "This means the disease-causing mutations previously reported are very rare, and thus probably first appeared relatively recently—that is, in the past several hundred years."

The results show how challenging it can be to interpret genetic data, adds Coller. "Some variants of these two genes will likely be obviously deleterious, but it may be impossible to predict whether others are deleterious," says Coller, also the David Rockefeller Professor in the Allen and Frances Adler Laboratory of Blood and Vascular Biology. "In those cases, we will need to use additional information to judge the likelihood of a mutation being deleterious, and in many cases there will

be residual uncertainty," Coller says.

If individuals have only one copy of a deleterious variant, they are unlikely to notice any impact on their health because there will be enough receptors to prevent bleeding. A more serious issue arises when two carriers of a mutated gene have children: If one percent of people carry a variant and if roughly half of the variants are deleterious, approximately three out of every 100,000 babies could carry two deleterious variants from their parents. With the human population continuing to expand at an unprecedented rate and new variants entering the genome with each generation, says Coller, it will become more important over time to use the powerful new genetic tools at our disposal to advise people of their potential genetic risks.

More information: Buitrago, L., Rendon, A., Liang, Y., Simeoni, I., Negri, A., Filizola, M., … Coller, B. S. (2015). αIIbβ3 variants defined by next-generation sequencing: Predicting variants likely to cause Glanzmann thrombasthenia. *Proceedings of the National Academy of Sciences*, 201422238. DOI: 10.1073/pnas.1422238112

Provided by Rockefeller University