

# IndiGenomes: a comprehensive resource of genetic variants from over 1000 Indian genomes

Abhinav Jain<sup>1,2,1</sup>, Rahul C. Bhojar<sup>1,1</sup>, Kavita Pandhare<sup>1</sup>, Anushree Mishra<sup>1</sup>, Disha Sharma<sup>1</sup>, Mohamed Imran<sup>1,2</sup>, Vigneshwar Senthivel<sup>1,2</sup>, Mohit Kumar Divakar<sup>1,2</sup>, Mercy Rophina<sup>1,2</sup>, Bani Jolly<sup>1,2</sup>, Arushi Batra<sup>1,2</sup>, Sumit Sharma<sup>1</sup>, Sanjay Siwach<sup>1</sup>, Arun G. Jadhao<sup>3</sup>, Nikhil V. Palande<sup>4</sup>, Ganga Nath Jha<sup>5</sup>, Nishat Ashrafi<sup>5</sup>, Prashant Kumar Mishra<sup>6</sup>, Vidhya A. K.<sup>7</sup>, Suman Jain<sup>8</sup>, Debasis Dash<sup>1,2</sup>, Nachimuthu Senthil Kumar<sup>9</sup>, Andrew Vanlalawma<sup>9</sup>, Ranjan Jyoti Sarma<sup>9</sup>, Lalchandama Chhakchhuak<sup>10</sup>, Shantaraman Kalyanaraman<sup>11</sup>, Radha Mahadevan<sup>11</sup>, Sunitha Kandasamy<sup>11</sup>, Pabitha B. M.<sup>11</sup>, Raskin Erusan Rajagopal<sup>11</sup>, Ezhil Ramya J.<sup>11</sup>, Nirmala Devi P.<sup>11</sup>, Anjali Bajaj<sup>1,2</sup>, Vishu Gupta<sup>1,2</sup>, Samatha Mathew<sup>1,2</sup>, Sangam Goswami<sup>1,2</sup>, Mohit Mangla<sup>1,2</sup>, Savinitha Prakash<sup>1</sup>, Kandarp Joshi<sup>1</sup>, Meyakumla<sup>1</sup>, Sreedevi S.<sup>12</sup>, Devarshi Gajjar<sup>13</sup>, Ronibala Soraislam<sup>14</sup>, Rohit Yadav<sup>1,2</sup>, Yumnam Silla Devi<sup>15</sup>, Aayush Gupta<sup>16</sup>, Mitali Mukerji<sup>1,2</sup>, Sivaprakash Ramalingam<sup>1,2</sup>, Binukumar B. K.<sup>1,2</sup>, Vinod Scarla<sup>1,2,\*</sup> and Sridhar Sivasubbu<sup>1,2,\*</sup>

<sup>1</sup>CSIR-Institute of Genomics and Integrative Biology, New Delhi 110025, India, <sup>2</sup>Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, Uttar Pradesh 201002, India, <sup>3</sup>Department of Zoology, RTM Nagpur University, Nagpur, Maharashtra 440033, India, <sup>4</sup>Department of Zoology, Shri Mathuradas Mohota College of Science, Nagpur, Maharashtra 440009, India, <sup>5</sup>Department of Anthropology, Vinoba Bhave University, Hazaribag, Jharkhand 825301, India, <sup>6</sup>Department of Biotechnology, Vinoba Bhave University, Hazaribag, Jharkhand 825301, India, <sup>7</sup>Department of Biochemistry, Dr. Kongu Science and Art College, Erode, Tamil Nadu 638107, India, <sup>8</sup>Thalassemia and Sickle cell Society, Hyderabad, Telangana 500052, India, <sup>9</sup>Department of Biotechnology, Mizoram University, Aizawl, Mizoram 796004, India, <sup>10</sup>Department of Pathology, Civil Hospital Aizawl, Mizoram 796001, India, <sup>11</sup>TVMC, Tirunelveli Medical College, Tirunelveli, Tamil Nadu 627011, India, <sup>12</sup>Department of Microbiology, St.Pious X Degree & PG College for Women, Hyderabad, Telangana 500076, India, <sup>13</sup>Department of Microbiology, The Maharaja Sayajirao University of Baroda, Vadodara, Gujarat 390002, India, <sup>14</sup>Department of Dermatology, Venereology and Leprology, Regional Institute of Medical Sciences, Imphal, Manipur 795004, India, <sup>15</sup>CSIR- North East Institute of Science and Technology, Jorhat, Assam 785006, India and <sup>16</sup>Department of Dermatology, Dr. D.Y. Patil Medical College, Pune, Maharashtra 411018, India

Received August 10, 2020; Revised October 01, 2020; Editorial Decision October 02, 2020; Accepted October 22, 2020

## ABSTRACT

With the advent of next-generation sequencing, large-scale initiatives for mining whole genomes and exomes have been employed to better understand global or population-level genetic architecture. India encompasses more than 17% of the world population with extensive genetic diversity, but is under-represented in the global sequencing datasets. This

gave us the impetus to perform and analyze the whole genome sequencing of 1029 healthy Indian individuals under the pilot phase of the 'IndiGen' program. We generated a compendium of 55,898,122 single allelic genetic variants from geographically distinct Indian genomes and calculated the allele frequency, allele count, allele number, along with the number of heterozygous or homozygous individuals. In the present study, these variants were sys-

\*To whom correspondence should be addressed. Tel: +91 11 29879109; Email: vinodsc@igib.in

Correspondence may also be addressed to Sridhar Sivasubbu, Tel: +91 11 29879106; Email: sridhar@igib.in; ssivasubbu@igib.res.in

<sup>†</sup>The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.