# IVaCS: an Integrated Variant Calling System

**Chiara M.[2], Chillemi G.[1], D'Antonio M.[1] , D'Onorio De Meo P.[1] , Flati T.[1,3], Gioiosa S.[1,3], Picardi E.[2,3], Pesole G.[2,3], Castrignanò T.[1]**

1. Cineca, Consorzio Interuniversitario di Supercalcolo, Bologna, Italy
2. Dipartimento di Bioscienze, Biotecnologie e Biofarmaceutica, Università degli Studi di Bari "A. Moro", Bari, Italy
3. Istituto di Biomembrane e Bioenergetica, Consiglio Nazionale delle Ricerche, Bari, Italy
4. Dipartimento di Bioscienze, Università degli Studi di Milano, Milano, Italy

The reduction in sequencing costs associated with next generation sequencing technologies (NGS) has led to a rapid upsurge in the amount of genome re-sequencing data, paving the way for the advent of personalized genomics and precision medicine. Accurate genotyping is crucial for effective analyses of these data, and in particular for the correct identification of candidate causal mutations in diagnostic screenings. The body of genome resequencing data will likely see exponential growth in the next few years, underlining the need for publicly available, accurate and time-effective bioinformatics systems for data analysis. Ideally, such systems should be easy to use and constantly updated as new genomes and software tools are released.

Here we present IVaCS, a fully automated, highly accurate system with a web based graphical interface for genotyping and variant annotation. IVaCS offers state of the art tools for variant calling and annotation along with expert made pipelines for the analysis of whole genome sequencing (WGS), whole-exome sequencing (WES) and targeted resequencing (TGS) data, performing all steps from quality trimming to variant annotation.

The system is specifically designed to assist users with little or no bioinformatics skills and all the pipelines are available through a user friendly web interface. The final output is provided in the form of a dynamic web page where variants can be selected on the base of user defined hard filters. A comprehensive report containing detailed information and statistics concerning the execution of each step of the pipelines is also generated.

Extensive tests on publicly available genome resequencing data (Illumina platinum genome NA12878), show that our system recovers a slightly better sensitivity and a higher specificity than the commercial Illumina VCAT 2.0 software. IVaCS is implemented with a modular architecture and each module (quality trimming, reads mapping, variant calling, variant annotation) can be used independently. IVaCS may manage all the major commercial kits for exome sequencing, such as Illumina, Agilent or Nimblegen, along with a comprehensive collection of reference genomes (all the Illumina genomes, including human, mouse and cow, among the others) with corresponding genomic annotations. Finally, the software leverages an ensemble of publicly available resources (e.g., dbSNP, OMIM, COSMIC and ClinVar among others) for the functional annotation of human variants. Advanced users needing more control over the single steps might also request the command-line version of the software which is more flexible and easy to customize.

IVaCS has a very active and growing community. The system is under constant development and new reference genomes, databases and bioinformatics tools are added to IvaCS on a regular basis.

IVACS is available at: **https://bioinformatics.cineca.it/ivacs**