**Workshop on Establishing a Central Resource
of Data from Genome Sequencing Projects**

Open Data Access Model

Wylie Burke, Lisa Brooks, Laura Lyman Rodriguez

In the open access model, data are made available for sharing by placing them on a publicly accessible web site. Data can be coded (so that readily identifiable data such as participants' names or addresses are not released) or they may be fully identified. The data envisioned for this model would include genomic (sequence, epigenomic, etc.), phenotype, and possibly environmental data for individual participants.

Advantages of the open access model

- Researchers (and others) have immediate access to data.
- Maximum potential for data use in terms of the research questions able to be addressed and the number and type of investigators able to ask them, e.g., trainees and senior investigators.

Disadvantages of the open access model

- No privacy protections are used beyond possibly coding. Although no system can guarantee privacy, public release of coded data does not provide the level of privacy protection as would public release of only summary data or a controlled-access process for researchers who agree to manage the data in certain ways to protect the data against privacy risks.
- Potential participants are likely to vary in their acceptance of this approach, and individuals from some populations or groups may be more willing to participate than others. Therefore, the resulting resource may not include representative samples. Such introduced bias may be difficult to measure or account for in study analyses.

Policy issues
- Informed consent processes for such a model would need to ensure participant awareness of the implications and potential privacy risks intrinsic to the open access framework. It would be important to develop (and understand how to best deploy) various approaches to convey the complex concepts important to decision making about open and full access to an individual's genomic information.
- If this model were to go forward, research seeking to understand factors (such as degree of trust in the research process) that influence willingness to participate in the open access model could help to define the potential for bias in this approach. Such research might also identify methods or initiatives to decrease sampling biases.