

---

## Plan Overview

*A Data Management Plan created using DMPonline*

**Title:** Technology development for cancer genome investigation

**Creator:** Afshin Ahmadian

**Principal Investigator:** Afshin Ahmadian

**Data Manager:** Afshin Ahmadian

**Affiliation:** KTH Royal Institute of Technology

**Template:** DCC Template

**ID:** 56500

**Last modified:** 22-04-2020

### **Copyright information:**

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

# Technology development for cancer genome investigation

---

## Data Collection

### What data will you collect or create?

What type, format and volume of data? TBs of sequencing data

Do your chosen formats and software enable sharing and long-term access to the data? Yes

Are there any existing data that you can reuse? No

### How will the data be collected or created?

Cancer patient materials that are sequenced for mutational analysis

What standards or methodologies will you use? NGS

How will you structure and name your folders and files? How will you handle versioning? Project number and project date and also version number

What quality assurance processes will you adopt? Sample replicates and peer review

## Documentation and Metadata

### What documentation and metadata will accompany the data?

Upon publication parts of the data from the study will be uploaded to the Sequence Read Archive (SRA) with an accompanying accession number

## Ethics and Legal Compliance

### How will you manage any ethical issues?

The project has ethical permission with the permission numbers 2015/2139-52, 2016/957-31

### How will you manage copyright and Intellectual Property Rights (IPR) issues?

When the study is published the authors will be the owner and have the copyright

## **Storage and Backup**

### **How will the data be stored and backed up during the research?**

We will store the sequencing data on SNIC Bianca. The data will be backed up and recovered in the event of incident by UppMax

### **How will you manage access and security?**

The data security will be managed by UppMax/Bianca while the collaborators may get access through login credentials.

## **Selection and Preservation**

### **Which data are of long-term value and should be retained, shared, and/or preserved?**

The data will be retained on UppMax for at least one year with possibility for extension

### **What is the long-term preservation plan for the dataset?**

Initially on UppMax/Bianca and in future using the long term storage on UppMax with a price of 800 SEK/TB and year

## **Data Sharing**

### **How will you share the data?**

Parts of the data will be published in peer reviewed journals

### **Are any restrictions on data sharing required?**

Yes, the whole sequencing data cannot be made available publicly.

## **Responsibilities and Resources**

**Who will be responsible for data management?**

The principle investigator is responsible for data management

**What resources will you require to deliver your plan?**

For long term storage small funds is needed