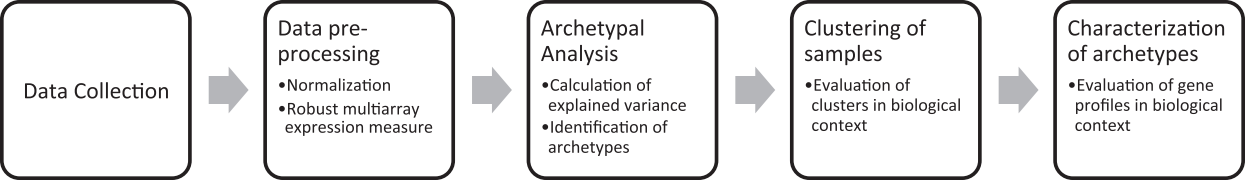


Data Collection



```
graph LR; A[Data Collection] --> B[Data pre-processing<br/>• Normalization<br/>• Robust multiarray expression measure]; B --> C[Archetypal Analysis<br/>• Calculation of explained variance<br/>• Identification of archetypes]; C --> D[Clustering of samples<br/>• Evaluation of clusters in biological context]; D --> E[Characterization of archetypes<br/>• Evaluation of gene profiles in biological context];
```

The diagram illustrates a five-step data analysis pipeline. It begins with 'Data Collection', followed by 'Data pre-processing' which includes normalization and robust multiarray expression measures. The third step is 'Archetypal Analysis', involving the calculation of explained variance and identification of archetypes. This is followed by 'Clustering of samples', which evaluates clusters in a biological context. The final step is 'Characterization of archetypes', which evaluates gene profiles in a biological context. Each step is contained within a rounded rectangular box, and the steps are connected by large grey arrows pointing from left to right.

Data pre-
processing

- Normalization
- Robust multiarray expression measure

Archetypal
Analysis

- Calculation of explained variance
- Identification of archetypes

Clustering of
samples

- Evaluation of clusters in biological context

Characterization
of archetypes

- Evaluation of gene profiles in biological context