

Manchot-Forschungsgruppe KI – Use-Case Gesundheit

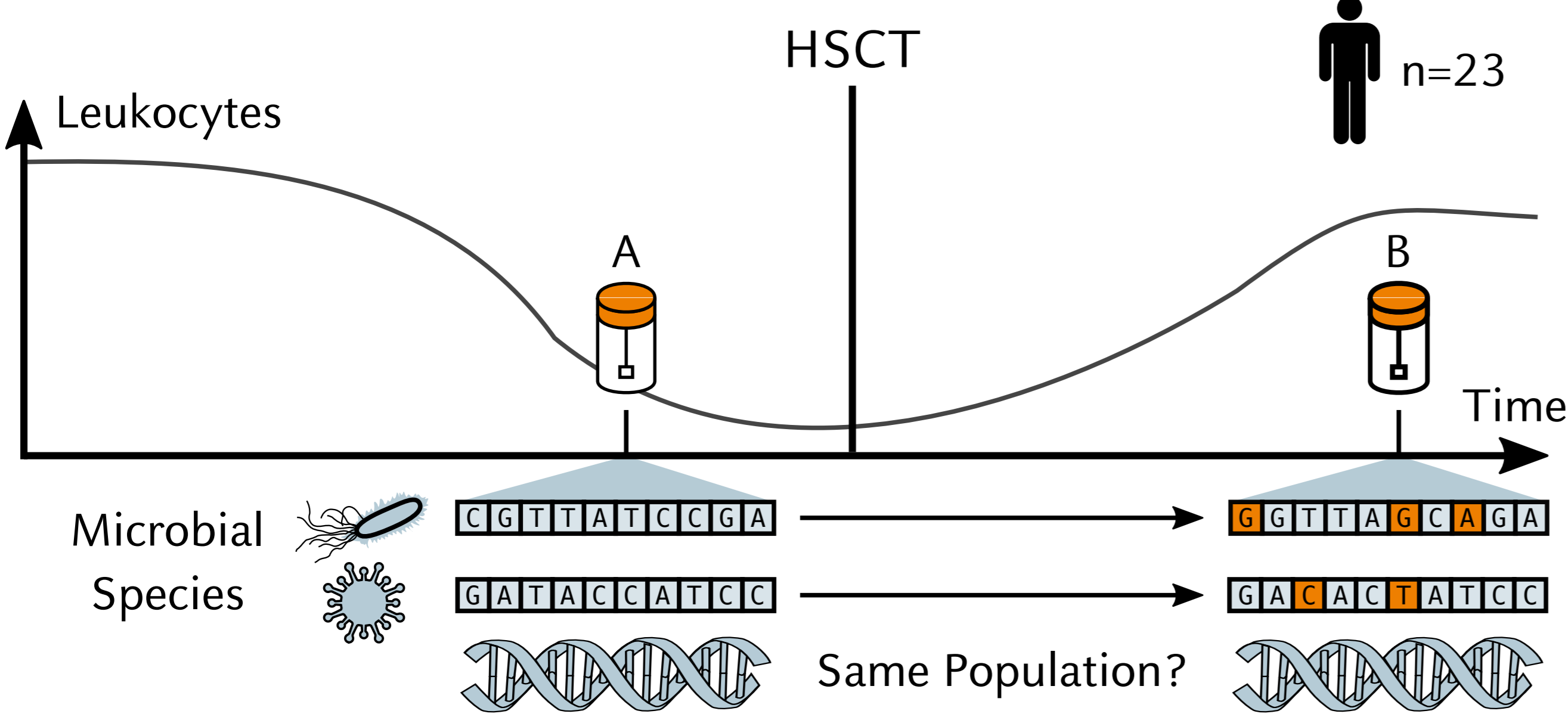
GutTrSnp: Detecting Population Shifts

Philipp Spohr¹ Anna Rommerskirchen² Sebastian Scharf² Paul Jäger³ Tobias Wienemann² Birgit Henrich² Gunnar Klau¹ Alexander Dilthey² Rainer Haas² Klaus Pfeffer²
¹Institute for Computer Science, HHU ²Institute of Medical Microbiology and Hospital Hygiene, UKD ³Department of Haematology, Oncology and Clinical Immunology, UKD

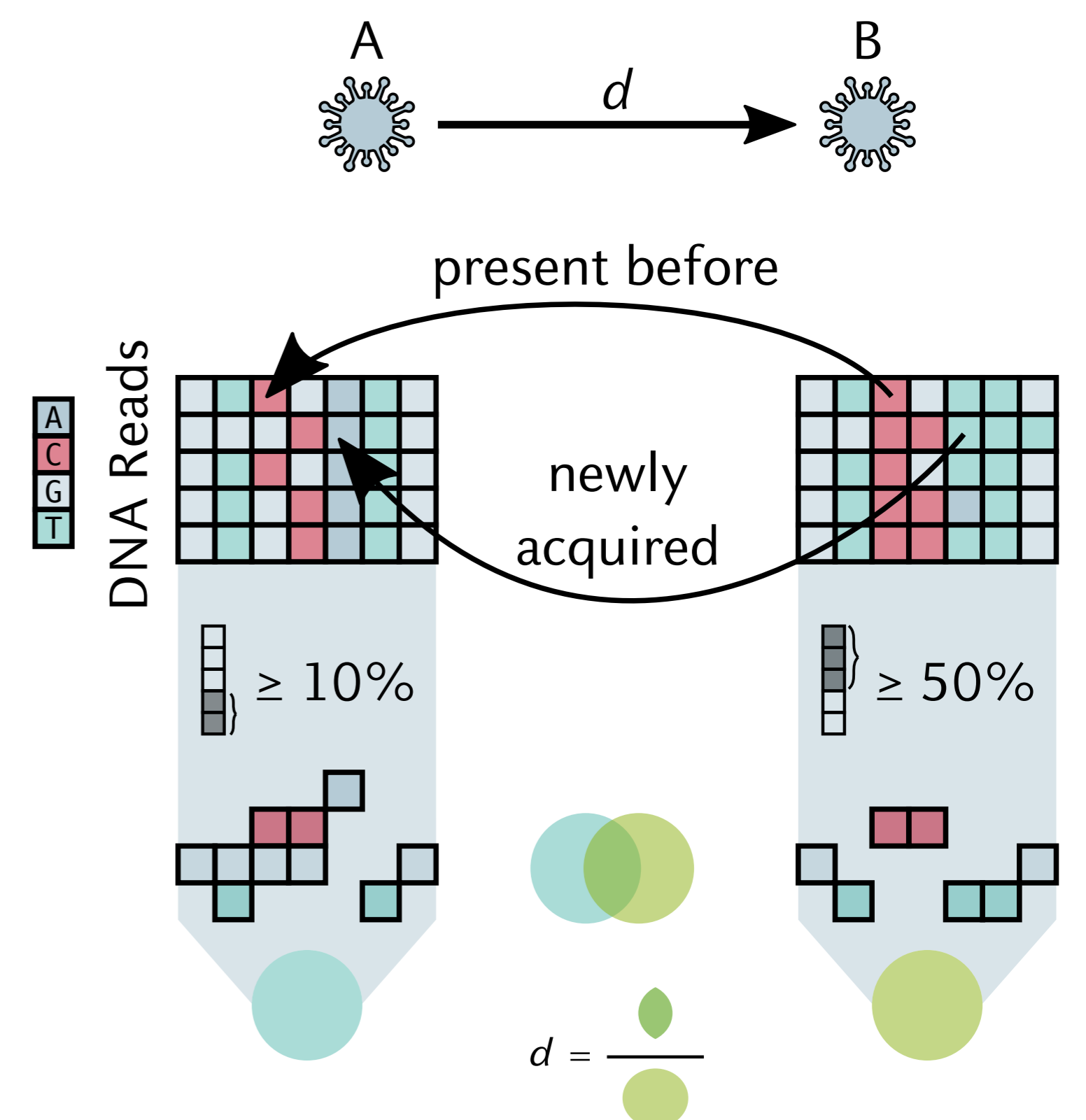
Abstract: Patients undergoing Haematopoietic Stem Cell Transplantation (HSCT) experience a massive disruption in their gut microbiome. One question of interest is whether species, observed before and after HSCT, survived or are the result of recolonization from an external source. To address this question, we developed a workflow that:

1. Calculates genome distances for each species and each pair of samples
2. Learns the genome substructure of each species based on the distances
3. Calculates likelihoods for each pair of samples being in distinct genome clusters

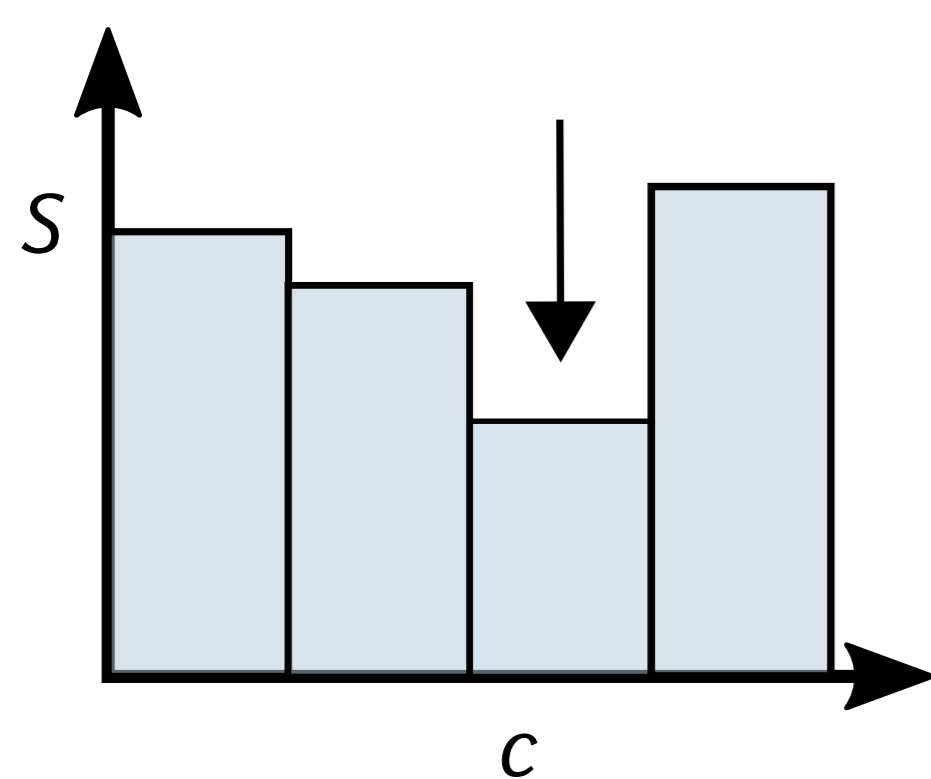
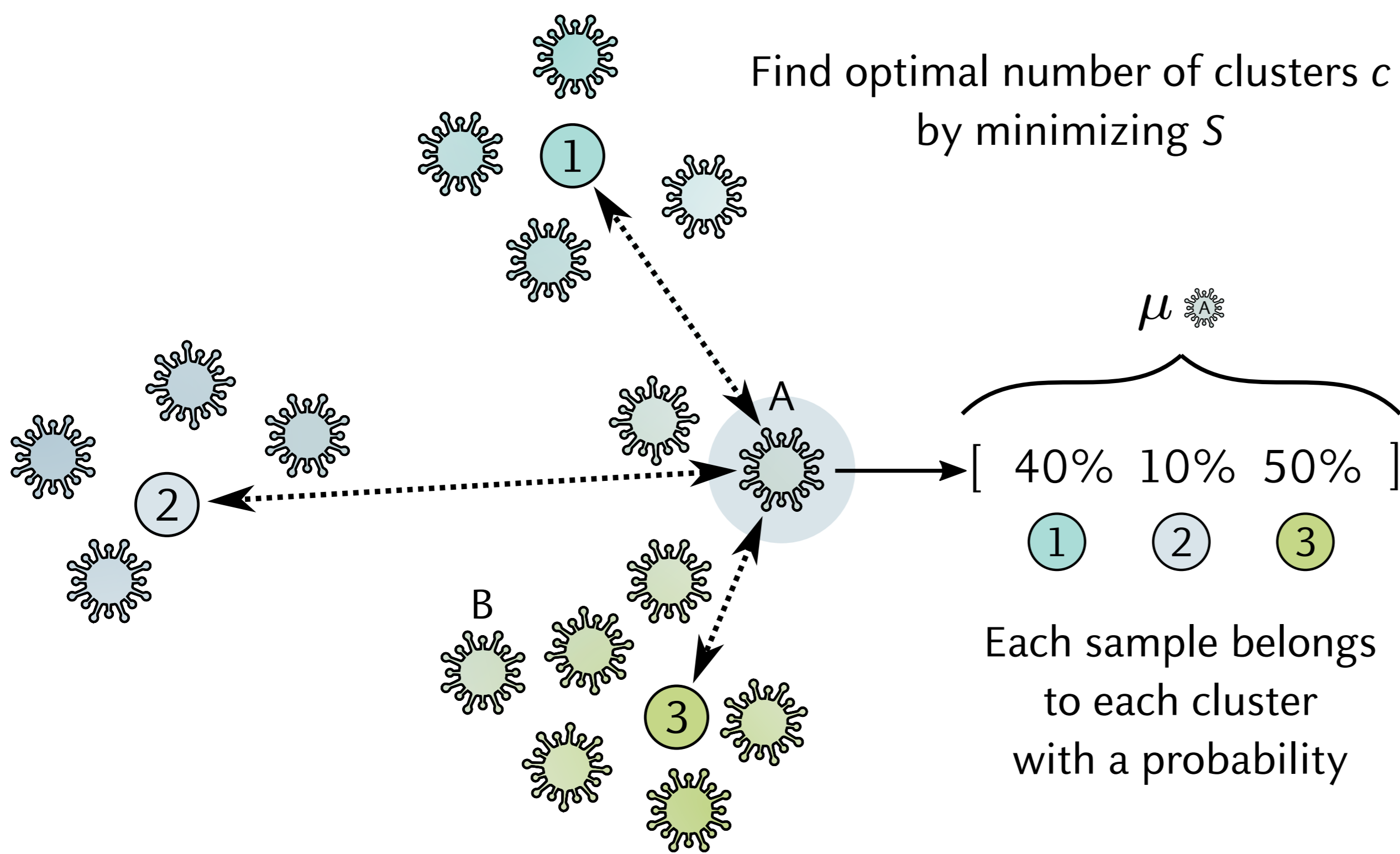
Scientific Question



1. Quantify Sample Distance Per Taxon



2. Learn Species Structure / Fuzzy Clustering



$$S = \frac{\sum_x \sum_{\mu_x} \mu_x^2 |x \leftrightarrow \mu_x|^2}{\sum_x \sum_y \min_{\mu_x, \mu_y} |x \leftrightarrow y|^2}$$

3. Classify Pairs as Population Shifts

