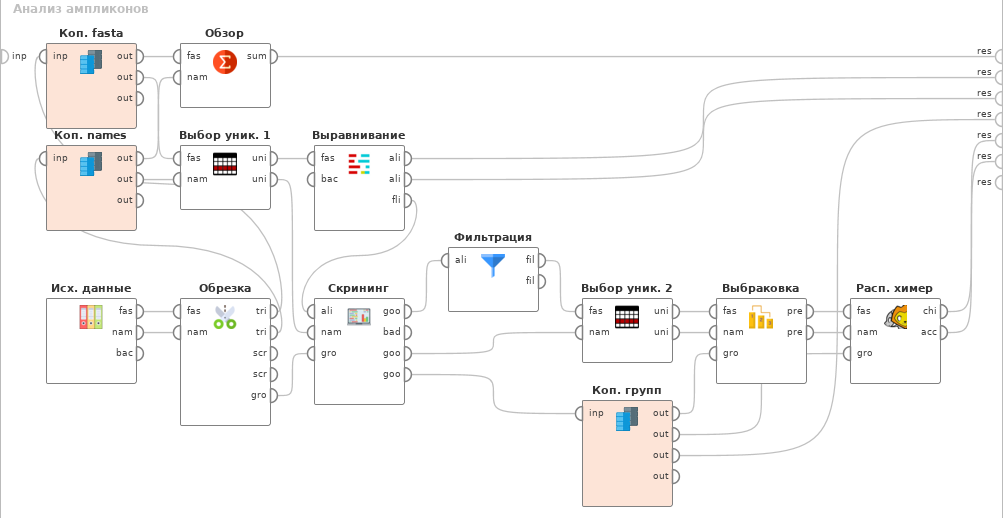
Предложена методика представления процесса обработки данных в анализе ампликонов в виде графов конвейерной обработки (Workflow). В соответствии с предложенной методикой реализован модуль расширения “New Generation Sequencing” для платформы Rapidminer, обеспечивающий композицию операций библиотеки Mothur в вычислительную процедуру (Рис. 1). Реализованы операции Trim, Summary, Unique, Align, Screen и др. Интеграция с платформой Rapidminer позволяет расширить существующий инструментарий исследований модулями платформы, в том силе модулями визуализации данных, стандартизировать процесс анализа ампликонов на основе современных методов организации вычислительных процессов в области Больших данных.



***Рис. 1.*** Сетевой граф начальной стадии процесса анализа ампликонов

Обобщение опыта реализации операций Mothur позволило перейти к разработке системы автоматизации, обеспечивающей генерацию (автоматичнское порождение) исходного кода процедур модуля расширения на основе анализа файлов журналирования программы Mothur, – дальнейшего развития проекта в направлении совершенствования информационно-вычислительных средств анализа.

1. Морозов А.А., Шигаров А.О., Малков Ф.С., Паскал К.К., Черкашин Е.А., Лихошвай Е.В. Информационно-вычислительная инфраструктура для поддержки метагеномного анализа // Вестник ИГУ. – 2017. (статья РИНЦ, в печати)
2. Cherkashin E., Shigarov A., Malkov F., Pascal K., Morozov A. An Environment for Metagenomic Analysis // In Proc. Int. Conf. on Applied Internet and Information Technologies. Zrenjanin, Serbia, 2017. pp. 110-117. (Доклад в трудах международной конференции)