Reference Assisted Assembly and Annotation of the Octopus vulgaris Genome GENOMICS CORELEUVEN **KU LEUVEN** Koen Herten, Gregory Maes, Eve Seuntjes, Fiorito Graziano and Joris R Vermeesch Thinking: Ivybridge Sequencing: 20 cores/node Octopus 213Gb (A,C,G,T) **124GB RAM** +90x coverage vulgaris HiSeq1500 PE95bp Cerebro: Ivybridge Estimated genome size: 10 cores/node Multiple insert sizes 2.4 - 4.6Gb 250GB shared Estimated # genes: 33,000 Reduction & De novo assembly Pre-computation 165 350 403 Scaffolding 852 GB GB GB GB Genome size: 2.36Gb Estimated genome size: Genome size: 3.39Gb 2.3Gb #Scaffolds: 7,452k #Scaffolds: 1,290k High Heterozygosity! N50: 3k N50: 3.5k cput 65h, 105GB RAM cput 2638h, 351GB RAM cput 470h, 293GB RAM **Annotation** Repeat Masking Reference Assisted 852 Scaffolding 857 861 856 GB GB GB GB Found 15,036 possible gene Genome size: 1.85Gb No stats yet #Scaffolds: 84k locations. N50: 276k cput 1.5h, 16GB RAM cput 50m, 2GB RAM cput 50m, 2.2GB RAM