A Big Earth Data Platform for Three Poles

**Genomic studies of drought tolerance mechanisms of a typical plant in Heihe basin - dataset II (2014-2015)**

1、Description

一. Data overview  
 This data interchange is the second data interchange of "genomics research on drought tolerance mechanism of typical desert plants in heihe basin", a key project of the major research program of "integrated research on eco-hydrological processes in heihe basin".The main research goal of this project is a typical desert sand Holly plants as materials, using the current international advanced a new generation of gene sequencing technology to the whole genome sequence and gene transcription of Holly group sequence decoding, so as to explore related to drought resistance gene and gene groups, and transgenic technology in model plants such as arabidopsis and rice) verify its drought resistance.  
二, data content  
1.Sequencing of the genome and transcriptome of lycophylla SPP.  
 The genome size of Mongolian Holly was about 926 Mb, GC content 36.88%, repeat sequence proportion 66%, genome heterozygosity rate 0.56%, which indicated that the genome has many repeat sequences, high heterozygosity and belongs to a complex genome.Based on the predicted sequence results, we subsequently carried out in-depth sequencing of the genome of lysiopsis SPP. The obtained data were assembled to obtain a 937 Mb genome sequence (table 1), which was basically the same as the predicted genome size.Through to the sand Holly transcriptome sequencing and sequence assembly (table 2), received more than 77000 genes coding sequence (Unigene), these sequences are comments found that most of the gene sequence and legumes and soybean, garbanzo beans and bean has a higher similarity (figure 1), consistent with the fact of sand ilex leguminous plants.  
一), and the sand Holly is a leguminous plants consistent with the fact.  
2.Discovery of simple repeat sequence (SSR) molecular markers of sand Holly:  
There is a transcriptome data set of sand Holly in the network public database, and the sample collection site is zhongwei city, ningxia.But this is the location of the project team samples in minqin county, gansu province, in order to study whether this sand in different areas of the Holly sequence has sequence polymorphism, we first identify the minqin county plant samples in the genomes of simple sequence repeat (SSR) markers (table 3), and then, compares the transcriptome sequences of plant sample, found in part of SSR molecular marker polymorphism (table 4), these molecular markers could be used for the species of plant genetic map construction, QTL mapping and genetic diversity analysis in the study.  
三, data processing instructions  
Sample collection place: minqin county, gansu province, latitude and longitude: N38 ° 34 '25.93 "E103 ° 08' 36.77".Genome sequencing: a total of 8 genomic DNA libraries of different sizes were constructed and determined by Illumina HiSeq 2500 instrument.Transcriptome sequencing: a library of 24 transcriptome mrnas was constructed and determined by Illumina HiSeq 4000.  
四, the use of data and meaning  
We selected a typical desert plant as the research object, from the Angle of genomics, parse the desert plant genome and transcriptome sequences, excavated its precious drought-resistant gene resources, and to study their drought resistance mechanism of favorable sand Holly this ancient and important to the utilization of plant resources, as well as the heihe river basin of drought-resistant plant genetic breeding, ecological restoration and sustainable development.

2、Keywords

Theme：Desert,Vegetation,Genomics,Desert ecosystem  
Discipline：Terrestrial Surface  
Places：Heihe River Basin, Minqin, Alashan  
Time：2015

3、Data details

1.Scale：10000

2.Projection：4326

3.Filesize：1.0MB

4.Data format：EXCEL

4、Space scope

|  |  |  |
| --- | --- | --- |
| - | north：38.79 | - |
| west：102.5 | - | east：103.57 |
| - | south：37.87 | - |

5、Time frame:2014-01-12 17:00:00+00:00--2016-01-11 17:00:00+00:00

6、Reference method

References to data:

HE Junxian. Genomic studies of drought tolerance mechanisms of a typical plant in Heihe basin - dataset II (2014-2015). A Big Earth Data Platform for Three Poles, doi:10.3972/heihe.0034.2016.db2016

References to articles:

7、Supporting project information

8、Data resource provider

name: HE Junxian  
unit:   
email: jxhe@cuhk.edu.hk