

Building Annotation Packages with `pdInfoBuilder` for Use with the `oligo` Package

Benilton Carvalho

April 7, 2009

1 Introduction

The `oligo` package offers support to multiple types of microarrays produced by Affymetrix and NimbleGen. The package will successfully read in CEL (Affymetrix) and XYS (NimbleGen) files, as long as the associated annotation package is already installed on the user's system.

The user must note that the annotation packages built for the `affy` package are **not** compatible with `oligo`. To have an annotation package that is compatible with `oligo`, one must use the `pdInfoBuilder` package.

This document shows examples on how to create such annotation packages for different platforms. After the package is created, the user must install it and not just copy it to the library tree.

2 The General Strategy for Building Annotation Packages with `pdInfoBuilder`

Building annotation packages with `pdInfoBuilder` depends on the followings files:

- Array design file: CDF (Affymetrix Expression), NDF (NimbleGen), BPMAP (Affymetrix Tiling) or PGF+CLF (Affymetrix Exon ST or Gene ST);
- Positions file: POS (NimbleGen - Tiling);
- Template of intensity file: CEL (Affymetrix) or XYS (NimbleGen);
- Probe sequence file: TAB (Affymetrix Expression)
- Probeset annotation file: PROBESET.CSV (Affymetrix Exon/Gene)

3 Loading the package

```
R> library(pdInfoBuilder)
```

Type	Package
50K Xba	pd.mapping50k.xba240
50K Hind	pd.mapping50k.hind240
250K Sty	pd.mapping250k.sty
250K Nsp	pd.mapping250k.nsp
SNP 5.0	pd.genomewidesnp.5
SNP 6.0	pd.genomewidesnp.6

Table 1: List of packages for SNP chips

4 Affymetrix SNP Arrays

4.1 Human Affymetrix SNP Arrays

The annotation packages for any Human Affymetrix SNP chip is available from BioConductor. We recommend the use of these packages as they contain additional data required to successfully run CRLMM.

4.2 Generic Affymetrix SNP Arrays

We also provide generic builder for Affymetrix SNP arrays. The user can use these tools to build basic annotation packages for SNP chips, but these will not contain data required by CRLMM, therefore CRLMM will not be available for arrays whose annotation packages were built using the method below.

It may be the case that the scripts do not find the required columns in the annotation and/or sequence files. If this happens, column names will be suggested and the user is expected to appropriately rename (or add) column names to the files.

4.2.1 Genome-Wide Arrays

For genome-wide arrays that contain both SNP and CNV probes, the strategy below should be followed. In this particular example, we use the SNP 6.0 array as an example (the user must note that for this particular array we recommend to download the pre-made packages from the BioConductor website).

```
R> baseDir <- "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffySnp6"
R> (cdf <- list.files(baseDir, pattern = ".cdf",
  full.names = TRUE))

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffySnp6/GenomeWideSNP_6.cdf"

R> (probeseqFileSNP <- list.files(baseDir,
  pattern = "6.probe_tab", full.names = TRUE))

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffySnp6/GenomeWideSNP_6.probe_ta
```

```

R> (probeseqFileCNV <- list.files(baseDir,
  pattern = "6.CN_probe_tab", full.names = TRUE))

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffySnp6/GenomeWideSNP_6.CN_probe

R> (annotFileSNP <- list.files(baseDir,
  pattern = "6.na24.annot.csv",
  full.names = TRUE))

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffySnp6/GenomeWideSNP_6.na24.ann

R> (annotFileCNV <- list.files(baseDir,
  pattern = "6.cn.na24.annot.csv",
  full.names = TRUE))

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffySnp6/GenomeWideSNP_6.cn.na24.

R> seed <- new("AffySNPCNVDPInfoPkgSeed2",
  cdfFile = cdf, csvAnnoFile = annotFileSNP,
  csvAnnoFileCnv = annotFileCNV,
  csvSeqFile = probeseqFileSNP,
  csvSeqFileCnv = probeseqFileCNV,
  version = "0.0.2", license = "Artistic",
  author = "Benilton Carvalho",
  email = "bcarvalh@jhsp.edu",
  biocViews = "AnnotationData",
  genomebuild = "NCBI Build 36",
  organism = "Human", species = "Homo Sapiens")
R> makePdInfoPackage(seed, destDir = ".")

=====
Building annotation package for Affymetrix SNP/CNV Array
CDF.....: GenomeWideSNP_6.cdf
SNP Annotation: GenomeWideSNP_6.na24.annot.csv
CNV Annotation: GenomeWideSNP_6.cn.na24.annot.csv
SNP Sequence..: GenomeWideSNP_6.probe_tab
CNV Sequence..: GenomeWideSNP_6.CN_probe_tab
=====
Parsing file: GenomeWideSNP_6.cdf ... OK
Getting SNP probes... OK
Organizing PM probes for SNPs... OK
Getting SNP information... OK
Organizing PM probes for CNVs... OK
Getting sequences for SNPs... OK
Getting sequences for CNVs... OK
Merging sequence information for SNPs... OK
Merging sequence information for CNVs... OK

```

```

Creating Biostrings objects... OK
Parsing file: GenomeWideSNP_6.na24.annot.csv ... OK
Merging information... OK
Parsing file: GenomeWideSNP_6.cn.na24.annot.csv ... OK
Merging information... OK
Creating package in ./pd.genomewidesnp.6
Inserting 909622 rows into table "featureSet"... OK
Inserting 5742214 rows into table "pmfeature"... OK
Inserting 945826 rows into table "pmfeatureCNV"... OK
Counting rows in featureSet
Counting rows in pmfeature
Counting rows in pmfeatureCNV
Creating index idx_pmfsetid on pmfeature ... OK
Creating index idx_pmfid on pmfeature ... OK
Creating index idx_pmfsetidcnv on pmfeatureCNV ... OK
Creating index idx_pmfidcnv on pmfeatureCNV ... OK
Creating index idx_fsfsetid on featureSet ... OK
Saving XDataFrame object for SNPs / PM.
Saving XDataFrame object for CNV / PM.
Done.

```

4.2.2 SNP Arrays

For SNP arrays (like the Human 50K and 250K chips), the strategy below should be followed. In this particular example, we use the NSP array as an example (the user must note that for this particular array we recommend to download the pre-made packages from the BioConductor website).

```

R> baseDir <- "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffySnpNsp"
R> (cdf <- list.files(baseDir, pattern = ".cdf",
  full.names = TRUE))

```

```

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffySnpNsp/Mapping250K_Nsp.cdf"

```

```

R> (probeseqFileSNP <- list.files(baseDir,
  pattern = "probe_tab", full.names = TRUE))

```

```

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffySnpNsp/Mapping250K_Nsp.probe_

```

```

R> (annotFileSNP <- list.files(baseDir,
  pattern = "annot.csv", full.names = TRUE))

```

```

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffySnpNsp/Mapping250K_Nsp.na28.a

```

```

R> seed <- new("AffySNPPDInfoPkgSeed2",
  cdfFile = cdf, csvAnnoFile = annotFileSNP,
  csvSeqFile = probeseqFileSNP,
  version = "0.0.2", license = "Artistic",

```

```

    author = "Benilton Carvalho",
    email = "bcarvalh@jhsp.edu",
    biocViews = "AnnotationData",
    genomebuild = "NCBI Build 36",
    organism = "Human", species = "Homo Sapiens")
R> makePdInfoPackage(seed, destDir = ".")

=====
Building annotation package for Affymetrix SNP Array
CDF.....: Mapping250K_Nsp.cdf
SNP Annotation: Mapping250K_Nsp.na28.annot.csv
SNP Sequence..: Mapping250K_Nsp.probe_tab
=====
Parsing file: Mapping250K_Nsp.cdf ... OK
Getting SNP probes... OK
Organizing PM probes for SNPs... OK
Getting SNP information... OK
Getting sequences for SNPs... OK
Merging sequence information for SNPs... OK
Creating Biostrings objects... OK
Parsing file: Mapping250K_Nsp.na28.annot.csv ... OK
Merging information... OK
Creating package in ./pd.mapping250k.nsp
Inserting 262338 rows into table "featureSet"... OK
Inserting 3223280 rows into table "pmfeature"... OK
Counting rows in featureSet
Counting rows in pmfeature
Creating index idx_pmfsetid on pmfeature ... OK
Creating index idx_pmfid on pmfeature ... OK
Creating index idx_fsfsetid on featureSet ... OK
Saving XDataFrame object for SNPs / PM.
Done.

```

5 Affymetrix HT-HGU133

For this particular array, the user must have access to three components: CDF, CEL (which will provide information on the array geometry) and probe sequence file (TAB-delimited).

```

R> baseDir <- "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffyHTHGU133P"
R> (cdf <- list.files(baseDir, pattern = ".CDF",
    full.names = TRUE))

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffyHTHGU133P/HT_HG-U133_Plus_PM."
R> (cel <- list.files(baseDir, pattern = ".CEL",
    full.names = TRUE)[1])

```

```

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffyHTHGU133P/Human_PM_TestData.A
R> (tab <- list.files(baseDir, pattern = "_tab",
  full.names = TRUE))

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffyHTHGU133P/HT_HG-U133_Plus_PM.

R> seed <- new("AffyExpressionPDInfoPkgSeed",
  cdfFile = cdf, celFile = cel,
  tabSeqFile = tab, author = "Benilton Carvalho",
  email = "bcarvalh@jhspk.edu",
  biocViews = "AnnotationData",
  genomebuild = "NCBI Build 36",
  organism = "Human", species = "Homo Sapiens",
  url = "http://www.biostat.jhsph.edu/~bcarvalh")
R> makePdInfoPackage(seed, destDir = ".")

```

```

=====
Building annotation package for Affymetrix Expression array
CDF.....: HT_HG-U133_Plus_PM.CDF
CEL.....: Human_PM_TestData.A01.CEL
Sequence TAB-Delim: HT_HG-U133_Plus_PM.probe_tab
=====
Parsing file: HT_HG-U133_Plus_PM.CDF ... OK
Parsing file: Human_PM_TestData.A01.CEL ... OK
Parsing file: HT_HG-U133_Plus_PM.probe_tab ... OK
Getting information for featureSet table... OK
Getting information for pm/mm feature tables ... OK
Combining probe information with sequence information ... OK
Getting sequence information for AFFX probes ...OK
Creating package in ./pd.ht.hg.u133.plus.pm
Inserting 54715 rows into table "featureSet"... OK
Inserting 519517 rows into table "pmfeature"... OK
Inserting 180 rows into table "mmfeature"... OK
Inserting 16943 rows into table "bgfeature"... OK
Counting rows in bgfeature
Counting rows in featureSet
Counting rows in mmfeature
Counting rows in pmfeature
Creating index idx_bgfsetid on bgfeature ... OK
Creating index idx_bgfid on bgfeature ... OK
Creating index idx_pmfsetid on pmfeature ... OK
Creating index idx_pmfid on pmfeature ... OK
Creating index idx_fssetid on featureSet ... OK
Saving XDataFrame object for PM.
Saving XDataFrame object for BG.
Done.

```

6 Affymetrix Tiling Array

```
R> baseDir <- "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffyTiling"  
R> (bmap <- list.files(baseDir, pattern = ".bmap",  
    full.names = TRUE))
```

```
[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffyTiling/Hs35b_P02R_v01-3_NCBIV34.bmap"
```

```
R> (cel <- list.files(baseDir, pattern = ".CEL",  
    full.names = TRUE)[1])
```

```
[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffyTiling/GSM178873.CEL"
```

```
R> seed <- new("AffyTilingPDInfoPkgSeed",  
    bmapFile = bmap, celFile = cel,  
    author = "Benilton Carvalho",  
    email = "bcarvalh@jhsp.hj.edu",  
    biocViews = "AnnotationData",  
    genomebuild = "NCBI Build 34",  
    organism = "Human", species = "Homo Sapiens",  
    url = "http://www.biostat.jhsph.edu/~bcarvalh")  
R> makePdInfoPackage(seed, destDir = ".")
```

```
=====  
Building annotation package for Affymetrix Tiling array  
BMAP: Hs35b_P02R_v01-3_NCBIV34.bmap  
CEL..: GSM178873.CEL  
=====  
Parsing file: Hs35b_P02R_v01-3_NCBIV34.bmap ... OK  
Getting geometry from CEL file... OK  
Getting PMs...OK  
Getting MMs...OK  
Getting background probes...OK  
Getting sequences...OK  
Creating package in ./pd.hs35b.p02r.v01  
Inserting 7 rows into table "chrom_dict"... OK  
Inserting 6020293 rows into table "pmfeature"... OK  
Inserting 1774 rows into table "mmfeature"... OK  
Inserting 37687 rows into table "bgfeature"... OK  
Counting rows in bgfeature  
Counting rows in chrom_dict  
Counting rows in mmfeature  
Counting rows in pmfeature  
Creating index idx_bgfid on bgfeature ... OK  
Creating index idx_pmfid on pmfeature ... OK  
Saving XDataFrame object for PM.  
Saving XDataFrame object for BG.  
Done.
```

7 Affymetrix Exon ST Array

```
R> library(pdInfoBuilder)
R> baseDir <- "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffyExon"
R> (pgf <- list.files(baseDir, pattern = ".pgf",
  full.names = TRUE))

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffyExon/HuEx-1_0-st-v2.r2.pgf"

R> (clf <- list.files(baseDir, pattern = ".clf",
  full.names = TRUE))

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffyExon/HuEx-1_0-st-v2.r2.clf"

R> (prob <- list.files(baseDir, pattern = ".probeset.csv",
  full.names = TRUE))

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffyExon/HuEx-1_0-st-v2.na27.hg18"

R> seed <- new("AffyExonPDInfoPkgSeed",
  pgfFile = pgf, clfFile = clf,
  probeFile = prob, author = "Benilton Carvalho",
  email = "bcarvalh@jhspsh.edu",
  biocViews = "AnnotationData",
  genomebuild = "NCBI Build 36",
  organism = "Human", species = "Homo Sapiens",
  url = "http://www.biostat.jhsph.edu/~bcarvalh")
R> makePdInfoPackage(seed, destDir = ".")

=====
Building annotation package for Affymetrix Exon ST Array
PGF.....: HuEx-1_0-st-v2.r2.pgf
CLF.....: HuEx-1_0-st-v2.r2.clf
Probeset: HuEx-1_0-st-v2.na27.hg18.probeset.csv
=====
Parsing file: HuEx-1_0-st-v2.r2.pgf ... OK
Parsing file: HuEx-1_0-st-v2.r2.clf ... OK
Creating initial table for probes...OK
Creating dictionaries... OK
Parsing file: HuEx-1_0-st-v2.na27.hg18.probeset.csv ... OK
Creating probeset -> gene table... OK
Creating genes table... OK
Creating package in ./pd.huex.1.0.st.v2
Inserting 100 rows into table "chrom_dict"... OK
Inserting 5 rows into table "level_dict"... OK
Inserting 8 rows into table "type_dict"... OK
Inserting 1625370 rows into table "fset2gene"... OK
Inserting 114281 rows into table "gene"... OK
```



```

Inserting 1425647 rows into table "featureSet"... OK
Inserting 5344479 rows into table "pmfeature"... OK
Inserting 37687 rows into table "bgfeature"... OK
Counting rows in bgfeature
Counting rows in chrom_dict
Counting rows in featureSet
Counting rows in fset2gene
Counting rows in gene
Counting rows in level_dict
Counting rows in pmfeature
Counting rows in type_dict
Creating index idx_bgfsetid on bgfeature ... OK
Creating index idx_bgfid on bgfeature ... OK
Creating index idx_pmfieldsetid on pmfeature ... OK
Creating index idx_pmfid on pmfeature ... OK
Creating index idx_fsfieldsetid on featureSet ... OK
Creating index idx_fs2gfsetid on fset2gene ... OK
Creating index idx_fs2ggid on fset2gene ... OK
Creating index idx_genegid on gene ... OK
Saving XDataFrame object for PM.
Saving XDataFrame object for BG.
Done.

```

8 Affymetrix Gene ST Array

```

R> library(pdInfoBuilder)
R> baseDir <- "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffyGene"
R> (pgf <- list.files(baseDir, pattern = ".pgf",
  full.names = TRUE))

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffyGene/HuGene-1_0-st-v1.r4.pgf"

R> (clf <- list.files(baseDir, pattern = ".clf",
  full.names = TRUE))

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffyGene/HuGene-1_0-st-v1.r4.clf"

R> (prob <- list.files(baseDir, pattern = ".probeset.csv",
  full.names = TRUE))

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/AffyGene/HuGene-1_0-st-v1.na27.2"

R> seed <- new("AffyGenePDInfoPkgSeed",
  pgfFile = pgf, clfFile = clf,
  probeFile = prob, author = "Benilton Carvalho",
  email = "bcarvalh@jhsph.edu",
  biocViews = "AnnotationData",

```

```

genomebuild = "NCBI Build 36",
organism = "Human", species = "Homo Sapiens",
url = "http://www.biostat.jhsph.edu/~bcarvalh")
R> makePdInfoPackage(seed, destDir = ".")

```

```

=====
Building annotation package for Affymetrix Gene ST Array
PGF.....: HuGene-1_0-st-v1.r4.pgf
CLF.....: HuGene-1_0-st-v1.r4.clf
Probeset: HuGene-1_0-st-v1.na27.2.hg18.probeset.csv
=====
Parsing file: HuGene-1_0-st-v1.r4.pgf ... OK
Parsing file: HuGene-1_0-st-v1.r4.clf ... OK
Creating initial table for probes...OK
Creating dictionaries... OK
Parsing file: HuGene-1_0-st-v1.na27.2.hg18.probeset.csv ... OK
Creating probeset -> gene table... OK
Creating genes table... OK
Creating package in ./pd.hugene.1.0.st.v1
Inserting 125 rows into table "chrom_dict"... OK
Inserting 5 rows into table "level_dict"... OK
Inserting 8 rows into table "type_dict"... OK
Inserting 1011778 rows into table "fset2gene"... OK
Inserting 87374 rows into table "gene"... OK
Inserting 257430 rows into table "featureSet"... OK
Inserting 764885 rows into table "pmfeature"... OK
Inserting 818005 rows into table "f2fset"... OK
Inserting 16943 rows into table "bgfeature"... OK
Counting rows in bgfeature
Counting rows in chrom_dict
Counting rows in f2fset
Counting rows in featureSet
Counting rows in fset2gene
Counting rows in gene
Counting rows in level_dict
Counting rows in pmfeature
Counting rows in type_dict
Creating index idx_bgfsetid on bgfeature ... OK
Creating index idx_bgfid on bgfeature ... OK
Creating index idx_pmfid on pmfeature ... OK
Creating index idx_f2fsfid on f2fset ... OK
Creating index idx_f2fsfsetid on f2fset ... OK
Creating index idx_fsfsetid on featureSet ... OK
Creating index idx_fs2gfsetid on fset2gene ... OK
Creating index idx_fs2ggid on fset2gene ... OK
Creating index idx_genegid on gene ... OK

```

```
Saving XDataFrame object for PM.  
Saving XDataFrame object for BG.  
Done.
```

9 NimbleGen Expression Array

```
R> library(pdInfoBuilder)  
R> baseDir <- "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/NgsExpression"  
R> (ndf <- list.files(baseDir, pattern = ".ndf",  
    full.names = TRUE))  
  
[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/NgsExpression/HG18_60mer_expr.ndf"  
  
R> (xys <- list.files(baseDir, pattern = ".xys",  
    full.names = TRUE)[1])  
  
[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/NgsExpression/9868701_532.xys"  
  
R> seed <- new("NgsExpressionPDInfoPkgSeed",  
    ndfFile = ndf, xysFile = xys,  
    author = "Benilton Carvalho",  
    email = "bcarvalh@jhspk.edu",  
    biocViews = "AnnotationData",  
    genomebuild = "NCBI Build 36",  
    organism = "Human", species = "Homo Sapiens",  
    url = "http://www.biostat.jhsph.edu/~bcarvalh")  
R> makePdInfoPackage(seed, destDir = ".")  
  
=====  
Building annotation package for Nimblegen Expression Array  
NDF:  HG18_60mer_expr.ndf  
XYS:  9868701_532.xys  
=====  
Parsing file: HG18_60mer_expr.ndf ... OK  
Parsing file: 9868701_532.xys ... OK  
Merging NDF and XYS files ...OK  
Preparing contents for featureSet table ...OK  
Preparing contents for bgfeature table ...OK  
Preparing contents for pmfeature table ...OK  
Creating package in ./pd.hg18.60mer.expr  
Inserting 24000 rows into table "featureSet"... OK  
Inserting 71998 rows into table "pmfeature"... OK  
Counting rows in featureSet  
Counting rows in pmfeature  
Creating index idx_pmfsetid on pmfeature ... OK  
Creating index idx_pmfid on pmfeature ... OK
```

```
Creating index idx_fsfsetid on featureSet ... OK
Saving XDataFrame object for PM.
Done.
```

10 NimbleGen Tiling Array

```
R> library(pdInfoBuilder)
R> baseDir <- "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/NgsTiling"
R> (ndf <- list.files(baseDir, pattern = ".ndf",
  full.names = TRUE))

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/NgsTiling/2006-07-18_HG18_RefSeq_

R> (xys <- list.files(baseDir, pattern = ".xys",
  full.names = TRUE)[1])

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/NgsTiling/92204_532.xys"

R> (pos <- list.files(baseDir, pattern = ".pos",
  full.names = TRUE))

[1] "/thumper/ctsa/snpmicroarray/organizing/pdInfoVignette/NgsTiling/2006-07-18_HG18_RefSeq_

R> seed <- new("NgsTilingPDInfoPkgSeed",
  ndfFile = ndf, xysFile = xys,
  posFile = pos, author = "Benilton Carvalho",
  email = "bcarvalh@jhsph.edu",
  biocViews = "AnnotationData",
  genomebuild = "HG 18", organism = "Human",
  species = "Homo Sapiens", url = "http://www.biostat.jhsph.edu/~bcarvalh")
R> makePdInfoPackage(seed, destDir = ".")

=====
Building annotation package for Nimblegen Tiling Array
NDF: 2006-07-18_HG18_RefSeq_promoter.ndf
POS: 2006-07-18_HG18_RefSeq_promoter.pos
XYS: 92204_532.xys
=====
Parsing file: 2006-07-18_HG18_RefSeq_promoter.ndf ... OK
Parsing file: 2006-07-18_HG18_RefSeq_promoter.pos ... OK
Parsing file: 92204_532.xys ... OK
Creating package in ./pd.2006.07.18.hg18.refseq.promoter
Inserting 18029 rows into table "featureSet"... OK
Inserting 386230 rows into table "pmfeature"... OK
Inserting 3077 rows into table "bgfeature"... OK
Counting rows in bgfeature
Counting rows in featureSet
```

```
Counting rows in pmfeature
Creating index idx_bgfsetid on bgfeature ... OK
Creating index idx_bgfid on bgfeature ... OK
Creating index idx_pmfsetid on pmfeature ... OK
Creating index idx_pmfid on pmfeature ... OK
Creating index idx_fsfsetid on featureSet ... OK
Saving XDataFrame object for PM.
Saving XDataFrame object for BG.
Done.
```

11 Session Information

```
R> sessionInfo()
```

```
R version 2.9.0 Under development (unstable) (2009-02-08 r47879)
x86_64-unknown-linux-gnu
```

```
locale:
```

```
LC_CTYPE=en_US.iso885915;LC_NUMERIC=C;LC_TIME=en_US.iso885915;LC_COLLATE=en_US.iso885915;LC_
```

```
attached base packages:
```

```
[1] stats      graphics  grDevices  utils
[5] datasets  methods   base
```

```
other attached packages:
```

```
[1] pdInfoBuilder_1.7.30 oligo_1.7.36
[3] preprocessCore_1.5.3 oligoClasses_1.5.20
[5] affxparser_1.15.2    RSQLite_0.7-1
[7] DBI_0.2-4            Biobase_2.3.11
```

```
loaded via a namespace (and not attached):
```

```
[1] affyio_1.11.3      Biostrings_2.11.44
[3] IRanges_1.1.55     splines_2.9.0
```