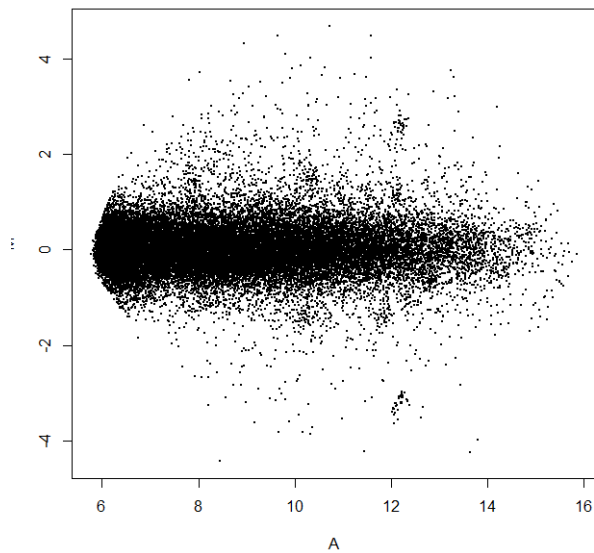


Zad 1

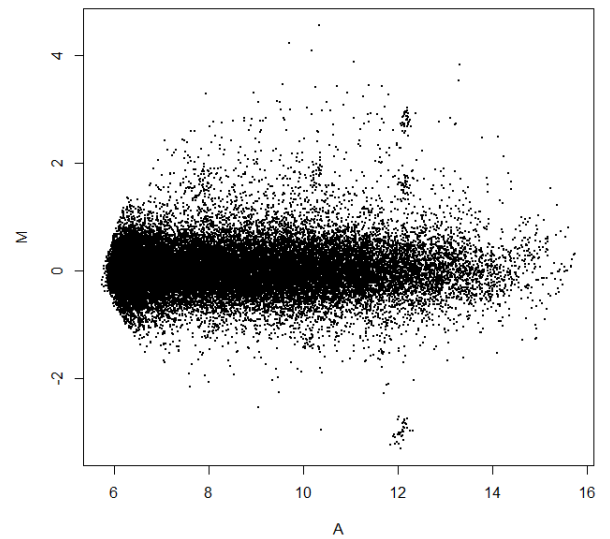
```
library(limma)

###ZAD1###
targets <- readTargets("C:/Users/marce/Desktop/biurowe 5/Analiza danych wysokoprzepust/z3/files/targets_agilent_cd.txt")
targets
RG <- read.maimages(targets, source="genepix.custom", path="C:/Users/marce/Desktop/biurowe 5/Analiza danych wysokoprzepust/z3/files/")
RG$genes=readGAL("C:/Users/marce/Desktop/biurowe 5/Analiza danych wysokoprzepust/z3/files/gal_do_anotacji2.gal")
RG$printer <- getLayout(RG$genes)
RG
RG <- backgroundCorrect(RG, method="normexp", offset=50)
MA <- normalizeWithinArrays(RG)
plotMD(MA[, 2])
RG
```

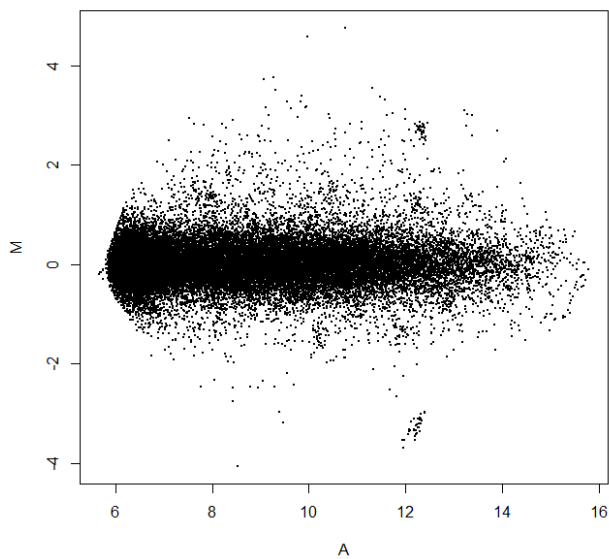
252111310223_morph31



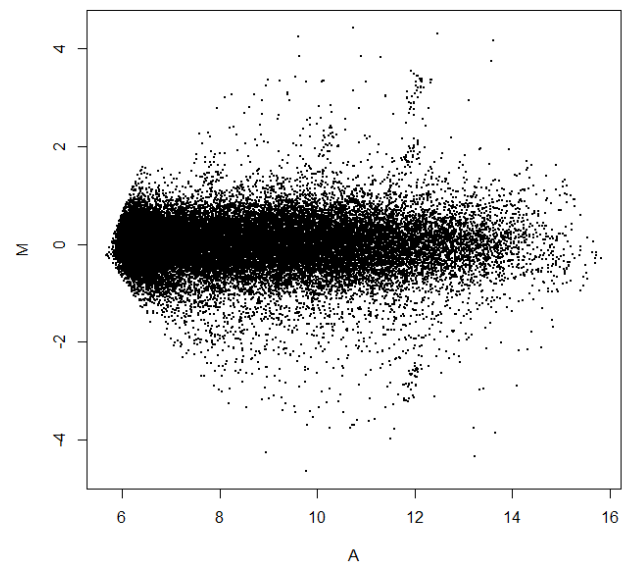
252111310223_morph32

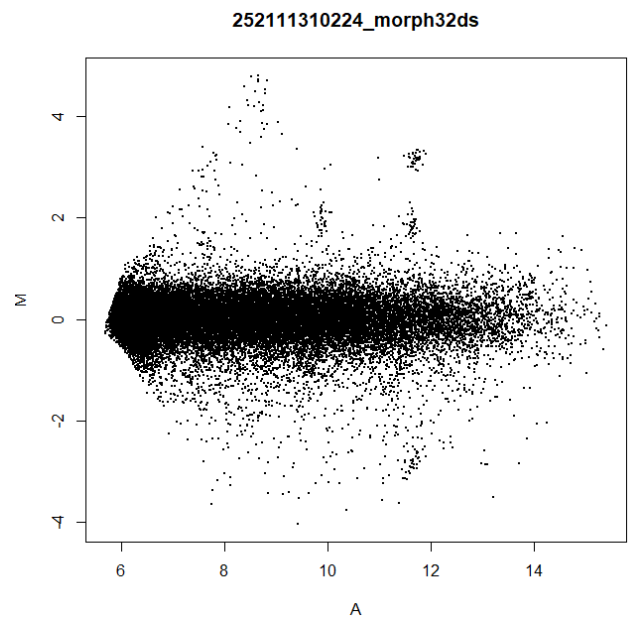
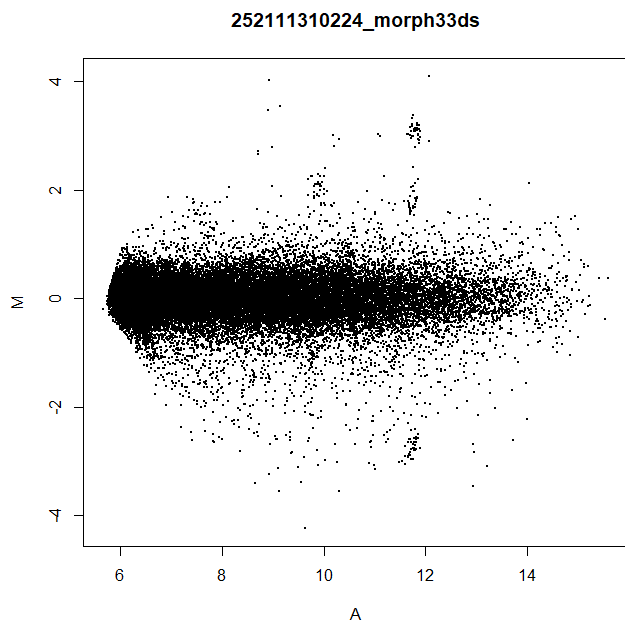


252111310223_morph33

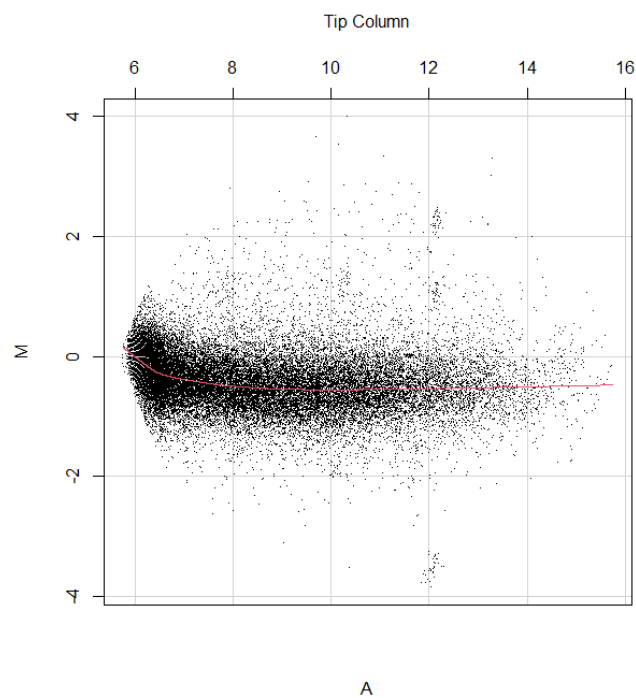


252111310223_morph31ds





Normalizacja przebiegła pomyślnie, ponieważ widać, że dla każdej macierzy punkty oscylują równo wokół krzywej wyznaczonej przez oś Y równą 0. Dla porównania wykres MAplot dla macierzy *morph32* przed normalizacją:



Widać tutaj, że te punkty nie układają się równomiernie przy osi y równej 0 i tworzy się tutaj krzywa bananowa.

Zad 2

```
###ZAD2###
design <- modelMatrix(targets, ref="Control")
design

fit <- lmFit(MA, design = design)
fit <- eBayes(fit)
typeof(fit)
#lista
topTable(fit, number=20)
#adjusted p-value -> jeśli niskie to różnicowa ekspresja
```

	CdCl2
252111310223_morph31	1
252111310223_morph32	1
252111310223_morph33	1
252111310223_morph31ds	-1
252111310224_morph32ds	-1
252111310224_morph33ds	-1

Obiekt powstały przy użyciu funkcji eBayes jest listą. Pole świadczące o różnicowej ekspresji to *adj. P value*.

	Block	Column	Row	Name	ID	RefNumber	ControlType	GeneName	TopHit	Description	logFC	AveExpr	t	P.Value	adj.P.Val	B
43603	1	245	164	AJ344597	A_95_P033279	3574	false	ta AJ344597			4.362177	9.689281	44.21195	4.691774e-13	2.121620e-08	19.01
33575	1	59	127	EB681340	A_95_P209072	35297	false	gb EB681340 ta TA19109_4097 tc TC10309		KL4B.113C20F.060131T7 KL48 Nicotiana tabacum cDNA db...	3.280540	8.980188	37.60455	2.449585e-12	5.538512e-08	17.92
26102	1	34	99	TA12394_4097	A_95_P013961	39561	false	ta TA12394_4097 tc TC6532		Q2VT56_CAPAN (Q2VT56) Probable glutathione-S-transferase...	2.924535	13.196244	33.39736	8.208082e-12	6.711790e-08	17.04
13044	1	10	50	TA15479_4097	A_95_P006331	43581	false	ta TA15479_4097 tc TC12039		Q39619_CATRO (Q39619) PAPS-reductase-like protein precu...	3.364856	9.677347	33.18427	8.760845e-12	6.711790e-08	16.99
8687	1	175	33	EB683065	A_95_P152722	15624	false	ta EB683065			2.641213	8.974474	33.07881	9.049380e-12	6.711790e-08	16.97
28640	1	178	108	EH619367	A_95_P273726	14992	false	gb EH619367 ta TA12340_4097		CHO_SL018xb09f1.ab1 CHO_SL Nicotiana tabacum cDNA. m...	3.510094	11.332844	32.38406	1.123187e-11	6.711790e-08	16.80
19932	1	248	75	EH615554	A_95_P157502	3193	false	ta EH615554			3.181566	9.369647	31.85319	1.328899e-11	6.711790e-08	16.67
19097	1	211	72	DQ460062	A_95_P215517	9400	false	gb DQ460062 ta TA20504_4097		Nicotiana tabacum cDNA-AFLP fragment H-N_BC1M33-1 se...	2.986417	9.603135	31.61340	1.435086e-11	6.711790e-08	16.61

Zad 3.

```
###ZAD3
t <- topTable(fit, number=45220) #that is number of all genes
t_p05 <- t[t['P.Value'] < 0.05,]
t_p01 <- t[t['P.Value'] < 0.01,]

results <- decideTests(fit)
summary(results)
vennDiagram(results)

volcanoplot(fit)
#najbardziej wysunięte do góry po bokach
```

10037 genów przeszło próg p-val 0.05 i 5961 próg 0.01.

t_p01	5961 obs. of 16 variables	
t_p05	10037 obs. of 16 variables	

```
> summary(results)
      CdCl2
Down    2231
NotSig 40143
Up      2846
```

Down oznacza obniżoną ekspresję a up podwyższoną. NotSig – brak różnic

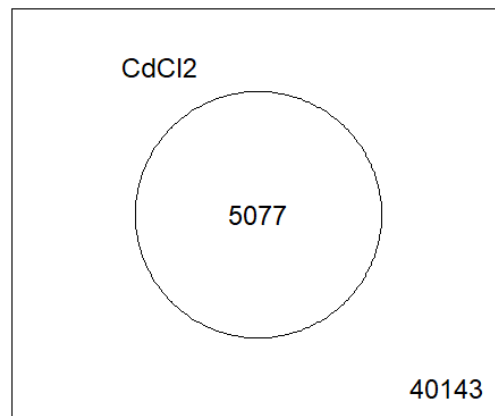
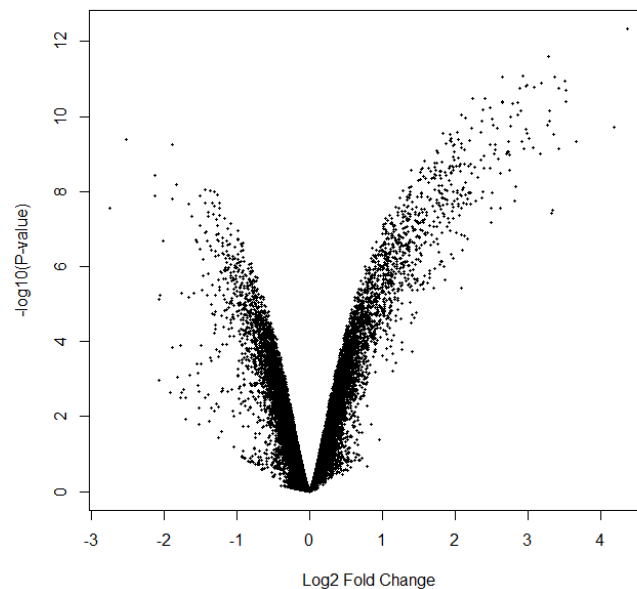


Diagram venna pokazuje ile genów sumarycznie posiada różnicową ekspresję. Jest to 5077 genów z 45220 wszystkich genów.



Volcano plot pokazuje istotność statystyczną (pvalue) w stosunku do siły zmiany (fold change). Podstawę / trzon wulkanu tworzą geny których ekspresja nie jest zróżnicowana – fold change w okolicach 0,1, -1. Również im większy logarytm ujemny z Pvalue (wyżej na wykresie) tym bardziej istotny statystycznie jest dana obserwacja. Geny najbardziej różnicowane są reprezentowane przez punkty wysunięte najbardziej na górnych bokach wykresu volcano plot.

Zad 4.

```
#zad4
tp <- topTable(fit, coef=1, number=Inf, adjust="BH")
top5 <- tp[tp['adj.P.Val'] < 0.05,]
top1 <- tp[tp['adj.P.Val'] < 0.01,]
```

top1	2859 obs. of 16 variables
------	---------------------------

top5	5077 obs. of 16 variables
------	---------------------------

```
top1_up <- top1[top1['logFC'] > 0,]
top1_down <- top1[top1['logFC'] < 0,]
write.csv(top1_up, 'up.csv')
write.csv(top1_down, 'down.csv')
```

Należy zauważyć, że genów tych będzie mniej niż w poprzednim zadaniu, gdyż tam brana pod uwagę była adj p value z progiem 0.05. Tutaj brany jest pod uwagę próg 0.01