

Supplementary Table 3. Estimates of population genetic parameters and LD in jack-knifed samples from 3 regions of *EGFR*.

Region ^a	N=12		N=48		N=96		Observed
		$\bar{X} \pm (\text{SD})$		$\bar{X} \pm (\text{SD})$		$\bar{X} \pm (\text{SD})$	
Population genetics							
1	Nuc. Differences ^b	7.75 (0.77)	7.75 (0.26)	7.73 (0.17)	7.73	(0.17)	7.73
	π	0.011 (0.001)	0.011 (0.0)	0.011 (0.0)	0.011	(0.0)	0.011
	π SD ^c	1.1E-03 (1.9E-04)	4.9E-04 (6.1E-05)	3.4E-04 (2.7E-05)	3.4E-04	(2.7E-05)	2.6E-05
	Sampling Var. π	1.2E-06 (4.3E-07)	2.4E-07 (6.4E-08)	1.1E-07 (2.2E-08)	1.1E-07	(2.2E-08)	1E-07
	q gene	7.03 (0.85)	6.70 (0.46)	6.79 (0.33)	6.79	(0.33)	7.24
	q site	0.0096 (0.0012)	0.0092 (0.0006)	0.0094 (0.0005)	0.0094	(0.0005)	0.0101
	Tajima D	0.47 (0.37)	0.54 (0.23)	0.44 (0.17)	0.44	(0.17)	0.20
	Fu Li D*	0.41 (0.37)	0.01 (0.49)	-0.75 (0.46)	-0.75	(0.46)	-1.71
	Fu Li F*	0.48 (0.40)	0.23 (0.46)	-0.34 (0.39)	-0.34	(0.39)	-1.08
	Haplotypes	10.11 (1.43)	29.58 (2.69)	47.39 (2.40)	47.39	(2.40)	68
	Haplotype Div.	0.96 (0.03)	0.96 (0.01)	0.96 (0.01)	0.96	(0.01)	0.96
	Hap. Div. SD ^c	0.05 (0.01)	0.02 (0.0)	0.01 (0.0)	0.01	(0.0)	0.01
	Seg. Sites	20.54 (2.44)	28.68 (2.00)	33.51 (1.56)	33.51	(1.56)	39
	Mutations	21.23 (2.57)	29.72 (2.03)	34.87 (1.71)	34.87	(1.71)	41
	Length	734.6 (5.5)	728.5 (6.8)	720.9 (5.6)	720.9	(5.6)	715
Linkage Disequilibrium							
	Nr. Tests	-	221.70 (35.30)	265.10 (29.20)	265.10	(29.20)	300
	Sites	-	21.50 (1.72)	23.50 (1.27)	23.50	(1.27)	25
	<0.05	-	29.6% (2.4%)	42.1% (3.3%)	42.1%	(3.3%)	49.3%
	<Bonferroni	-	8.5% (2.2%)	15.6% (2.4%)	15.6%	(2.4%)	22.3%
Population genetics							
2	Nuc. Differences ^b	9.40 (0.86)	9.32 (0.30)	9.31 (0.15)	9.31	(0.15)	9.30
	π	0.010 (0.001)	0.010 (0.0)	0.010 (0.0)	0.010	(0.0)	0.010
	π SD ^c	9.4E-04 (2.5E-04)	3.7E-04 (9.2E-05)	2.6E-04 (4.0E-05)	2.6E-04	(4.0E-05)	2.2E-05
	Sampling Var. π	9.5E-07 (5.5E-07)	1.5E-07 (7.0E-08)	7.4E-08 (4.4E-08)	7.4E-08	(4.4E-08)	0
	q gene	8.28 (1.25)	7.69 (1.03)	8.22 (0.83)	8.22	(0.83)	8.71
	q site	0.0084 (0.0013)	0.0079 (0.0011)	0.0085 (0.0009)	0.0085	(0.0009)	0.0090
	Tajima D	0.66 (0.48)	0.77 (0.44)	0.45 (0.35)	0.45	(0.35)	0.21
	Fu Li D*	0.50 (0.46)	-0.18 (0.75)	-1.54 (0.74)	-1.54	(0.74)	-2.60
	Fu Li F*	0.61 (0.52)	0.19 (0.75)	-0.89 (0.70)	-0.89	(0.70)	-1.70
	Haplotypes	9.62 (1.19)	27.19 (2.48)	43.72 (2.03)	43.72	(2.03)	55
	Haplotype Div.	0.95 (0.03)	0.94 (0.01)	0.94 (0.01)	0.94	(0.01)	0.94
	Hap. Div. SD ^c	0.05 (0.01)	0.02 (0.00)	0.01 (0.00)	0.01	(0.00)	0.01
	Seg. Sites	24.51 (3.59)	33.17 (4.54)	41.24 (4.26)	41.24	(4.26)	47
	Mutations	25.00 (3.77)	34.12 (4.58)	42.24 (4.26)	42.24	(4.26)	48
	Length	988.9 (8.1)	976.9 (9.3)	965.8 (6.2)	965.8	(6.2)	961
Linkage Disequilibrium							
	Nr. Tests	-	267.70 (35.57)	302.70 (18.12)	302.70	(18.12)	351
	Sites	-	23.60 (1.51)	25.10 (0.74)	25.10	(0.74)	27
	<0.05	-	37.9% (3.1%)	52.0% (2.3%)	52.0%	(2.3%)	52.7%
	<Bonferroni	-	12.5% (1.4%)	23.6% (1.1%)	23.6%	(1.1%)	25.6%

Supplementary Table 3. (Continued)

		N=12	N=48	N=96	
	Population genetics				
3	Nuc. Differences ^b	9.85 (1.09)	9.71 (0.49)	9.47 (0.30)	9.27
	π	0.009 (0.001)	0.009 (0.0)	0.009 (0.0)	0.009
	π SD ^c	1.0E-03 (1.9E-04)	4.8E-04 (4.1E-05)	3.4E-04 (1.7E-05)	2.8E-05
	Sampling Var. π	1.1E-06 (4.0E-07)	2.4E-07 (4.8E-08)	1.0E-07 (1.4E-08)	1E-07
	q gene	8.79 (1.11)	7.82 (0.52)	7.28 (0.35)	7.26
	q site	0.0081 (0.0010)	0.0073 (0.0005)	0.0069 (0.0003)	0.0070
	Tajima D	0.56 (0.36)	0.83 (0.25)	0.94 (0.18)	0.84
	Fu Li D*	0.46 (0.44)	0.53 (0.42)	0.22 (0.37)	-0.49
	Fu Li F*	0.55 (0.46)	0.76 (0.41)	0.61 (0.34)	0.07
	Haplotypes	10.13 (1.47)	28.48 (2.26)	43.16 (2.39)	53
	Haplotype Div.	0.96 (0.03)	0.96 (0.01)	0.96 (0.00)	0.96
	Hap. Div. SD ^c	0.05 (0.01)	0.01 (0.0)	0.01 (0.00)	0.01
	Seg. Sites	26.54 (3.35)	34.48 (2.22)	36.74 (1.67)	39
	Mutations	26.55 (3.36)	34.69 (2.33)	37.41 (1.78)	40
	Length	1078.9 (10.1)	1068.9 (15.0)	1050.8 (14.0)	1039
	Linkage Disequilibrium				
	Nr. Tests	-	402.20 (58.79)	426.40 (19.79)	435
	Sites	-	28.80 (2.04)	29.70 (0.67)	30
	<0.05	-	29.2% (7.5%)	46.0% (2.7%)	52.9%
	<Bonferroni	-	6.1% (2.0%)	16.1% (1.3%)	23.2%

Population genetic parameters were calculated from 100 jackknife subsamples (N=12, 48 and 96) from original samples of size 156 for region 1 and 139 for regions 2 and 3 (which were used for the observed data). Linkage disequilibrium was only estimated on 10 subsets, excluding N=12 sets for statistical reasons (LEWONTIN 1995).

a. Region 1 corresponds to exon 1 and flanking regions; region 2 includes last 350 bp of exon 4, exon 5, introns 4 and 5 in entirety and the first 350 bp of exon 6; and region 3 spans 1083 bp of exon 6.

b. Average number of nucleotide differences.

c. SD standard deviation.